Negative supercoil at gene boundaries modulates gene topology

Negative supercoil at gene boundaries modulates gene topology

Yathish Jagadheesh Achar[1*], Mohamood Adhil[1], Ramveer Choudhary[1], Nick Gilbert[2] and Marco Foiani[1,3*].

[1] IFOM (Fondazione Istituto FIRC di Oncologia Molecolare) Via Adamello 16, 20139 Milan, Italy.
[2] Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK.
[3] Università degli Studi di Milano, Italy.

[*]Correspondence to: yathish.achar@ifom.eu or marco.foiani@ifom.eu

Transcription challenges the integrity of replicating chromosomes by generating topological stress and conflicts with forks[1,2]. Top1 and Top2 DNA topoisomerases and Hmo1, an HMGB family protein, assist DNA replication and transcription[3-6].We investigated the topological architecture of genes in G1 and S phase in *Saccharomyces cerevisiae*. We found under-wound DNA at gene boundaries and over-wound DNA within coding regions. This arrangement does not depend on Pol II or S phase. Top2 and Hmo1 preserve negative supercoil at gene boundaries, while Top1 acts at coding regions. Transcription generates RNA:DNA hybrids within coding regions, independently of fork orientation. In S phase, Hmo1 protects under-wound DNA from Top2, while Top2 confines Pol II and Top1 at coding units, counteracting transcription leakage and aberrant hybrids at gene boundaries. Negative supercoil at gene boundaries prevents supercoil diffusion and nucleosome repositioning at coding regions. DNA looping occurs at Top2 clusters. We propose that Hmo1 locks gene boundaries in a cruciform conformation and, with Top2, modulates the architecture of genes that retain the memory of the topological arrangements even when transcription is repressed.

RNA polymerases generate positive and negative supercoils ahead and behind transcription bubbles, respectively[7]. Positive supercoiling accumulates in front of replication forks, precatenanes generate behind forks[8,9]. TMP (4,5′,8-Trimethylpsoralen) has been used to map DNA supercoiling[10-13] as psoralen intercalation is proportional to the negative super helical tension[14]. Using biotinylated TMP (bTMP)[12], we investigated the topology of transcribed genes and the contribution of Top1, Top2 and Hmo1 in maintaining the topological architecture of transcription units.

Topological context of Pol II genes

We analysed Rpb3 (Pol II subunit), Top2, Top1 and Hmo1 distributions in S phase and performed a meta-analysis on Pol II-transcribed genes (Fig. 1a). Rpb3 accumulated at ORFs, peaking at initiation (TSS) and termination (TTS) sites, likely reflecting slow transcription modes where transcription begins, and where transcription-coupled transactions occurs at termination[15]. Top2 and Hmo1 accumulated upstream and downstream of ORFs. Top1 was confined within the ORFs, accumulating close to TTS.

Using bTMP we mapped negative and positive supercoil, and stable regions[12] (Extended Data Fig. 1a,b). The three topological clusters were distributed near-equally (Extended Data Fig 1c). Negative supercoiled regions were mainly found at intergenic regions (49%), transcribed units exhibited a positive supercoiled context (40%) (Fig. 1b,c). Nucleosome occupied regions were distributed near-equally within the three topological territories (Extended Data Fig. 1c). Negative supercoil mirrored Top2 clusters (Fisher exact test *p-value* <1e-5), positive supercoil reflected Top1 distribution (Fisher exact test *p-value* <1e-20). Gene topology was comparable in G1 and S (Fig. 1b,c and Extended Data Fig. 1d,e).

We analyzed the topological profiles of conditionally expressed genes such as *ASF2* (transcribed in S phase) and the galactose-inducible gene cluster (Extended Data Fig. 1f,g). We found comparable bTMP profiles at *ASF2* in G1 (repressing conditions) and S, and at the Gal genes in glucose (repressing conditions) and galactose. We analyzed the locus containing the highly expressed *LEU2* gene, and the two moderately expressed *NFS1* and *DCC1* genes (Extended Data Fig. 1h). Rpb3 accumulated at *LEU2* but was undetectable at *NFS1* and *DCC1*. However, Top1 accumulated at *LEU2,* as well as, at *NFS1* and *DCC1*. Top2 and Hmo1 were present at gene boundaries. Hence, the topological context of Pol II transcribed genes does not depend on Pol II and the negative supercoil context at gene boundaries does not depend on Top2, since Top2 is recruited after G1[16].

We compared supercoil distribution in high, medium and low gene expression classes (Extended Data Fig. 2a). Highly expressed genes accumulated more Top2 and negative supercoil, compared to the

other two classes (Fig 1d,e). Conversely, highly expressed genes exhibited less positive supercoil (Extended Data Fig. 2b). Pol II and Top1 accumulation mirrored the levels of expression (Fig. 1d). Hmo1 distribution was comparable in the three classes of expression (Fig. 1d). Hence, under-wound DNA distribution at Pol II gene boundaries is enhanced in highly expressed genes.

We compared the supercoil context at intergenic spaces with respect to gene orientation by grouping Pol II genes into codirectional (+ and – strands), converging and diverging classes (Fig. 1f). Intergenic spaces between converging genes were smaller compared to the other directional classes. Divergent genes exhibited larger intergenic spaces. Converging genes accumulated more positive supercoil at intergenic spaces at the expenses of negative supercoil (Fig. 1g). Accordingly, converging intergenic regions exhibited lower Top2 binding, whereas Top1 binding was not affected (Extended Data Fig. 2c). Hence, convergent and divergent transcription have imposed specific topological context at intergenic spaces.

Top2 and Hmo1 contribute to gene topology

Temperature sensitive *top2-1* mutants exhibited a reduction of negative supercoil at gene boundaries and an increase of positive supercoil at the same regions (Fig. 2a and Extended Data Fig. 2d,h). In G1 *top2-1* mutants did not cause negative supercoil reduction (Extended Data Fig. 2e). *top1Δ* cells did not affect the topological context of gene boundaries or transcribed regions (Fig. 2a and Extended Data Fig. 2f,h). *top1Δ top2-1* mutants, like *top2-1* exhibited a decrease of negative supercoil at gene boundaries (Fig. 2a), although the accumulation of positive supercoil at the same regions was lower compared to *top2-1* cells (Extended Data Fig. 2g); *top1Δ top2-1* mutants accumulated less positive supercoil at transcribed regions compared to wt cells, suggesting that during transcription, the two topoisomerases can substitute for each other in maintaining a positive supercoiled context. Top1 localization in wt cells was restricted to coding regions while, in *top2-1* mutants accumulated at gene boundaries (Fig. 2b and Extended Data Fig. 2i). Thus, Top2 restricts Top1 at transcribed regions and *top2* mutants exhibit the unscheduled relocation of Top1 at gene boundaries, that can account for the local increase of positive supercoil. Moreover, in *top2-1* mutants Pol II, accumulated more at gene boundaries (Fig. 2c and Extended Data Fig. 2i). Hence, Top2 confines the transcription apparatus within the coding regions.

hmo1Δ cells exhibited a reduction of negative supercoil and accumulation of positive supercoil at gene boundaries, resembling *top2-1* mutants (Fig. 2d and Extended Data Fig. 3a,d). *HMO1* ablation in *top2-1* mutants restored a wt-like topological context at transcribed genes (Fig. 2d and Extended Data Fig. 3b,d). Top1 distribution in *HMO1Δ* and *HMO1Δ top2-1* mutants were similar to the one of wt cells (Extended Data Fig. 3c,e). Thus, the gene topological profiles of *HMO1Δ* and *top2-1* are comparable, but *HMO1Δ*, differently from *top2-1* mutation, does not cause Top1 accumulation at gene boundaries. Possibly, in *top2-1* mutants, Top1 accumulation at gene boundaries depends on DNA substrates generated by Hmo1.

Top2 restricts RNA:DNA hybrid within ORFs

Using the SF9 antibody[17,18], we addressed whether RNA-DNA hybrid accumulation reflected a specific topological context. In wt cells, hybrids were distributed within ORFs peaking at TTS (Extended Data Fig. 4a). Their accumulation did not correlate with expression levels (Extended Data Fig. 4b). RNaseH and the Rrm3 and Sen1 helicases counteract hybrid accumulation[19]. In *rnh1Δ* mutant hybrids accumulated throughout Pol II gene units, in *rrm3Δ* and *sen1^{cl}* mutants at TTS sites (Extended Data Fig. 4c); this is consistent with Rrm3 function in dismantling RNA transcripts while traveling on the lagging strand[20] and with Sen1 role in facilitating transcription termination[21]. These observations suggest that RNA:DNA hybrids represent a physiological intermediate during transcription, they are confined within coding regions, and their accumulation close to TTS may reflect the slowdown of Pol II elongation at termination[22].

In *top2* mutants, specifically in S phase, hybrids accumulated at gene boundaries, where negative supercoil reduction occurs (Fig. 2e). *top1top2* double mutants, resembled *top2* mutants (Extended Data Fig. 4e). *top1Δ* cells accumulated hybrids throughout the gene bodies (Fig. 2e), perhaps due to frequent Pol II pausing and backtracking; since *top1Δ* viability depends on Top2, it is possible that *top1Δ* cells phenocopy a Top2 defect, leading to Pol II leakage and hybrid accumulation at gene boundaries. Previous findings implicated Top1 in preventing hybrid accumulation[23]. Hence, Top2 counteracts hybrid accumulation and, in *top2* mutants, the accumulation of hybrids at flanking regions reflects the local decrease of negative supercoil and aberrant Pol II transcription. *hmo1Δ* cells exhibited a drastic reduction of hybrid accumulation compared to wt cells, and *hmo1Δ top2-1* mutants behaved similar to *hmo1Δ* cells (Extended Data Fig. 4f,g). Hence, *HMO1* ablation also rescued the aberrant accumulation of hybrids at flanking regions in *top2* mutants.

We addressed whether a clash between forks and transcribed genes might influence the accumulation of hybrids by analysing 173 efficient replication origins[24]. Transcription units in a head-on or co-directional orientation with replication forks within 0.25, 0.5, 1, 2 or 5 kbs from the origin point were selected. There was a significant enrichment of transcribed genes oriented head-on with replication forks (Fig. 2f); this reflects the overlap between the signals specifying transcription termination and those promoting replication initiation[25]. However, the relative accumulation of hybrids in the head-on and the co-directional classes of genes were comparable (Fig. 2f and Extended Data Fig. 4h). Notably, the intergenic regions of converging genes were prone to accumulate hybrids while this was not the case for the intergenic regions of divergent genes (Extended Data Fig. 4i).

Negative supercoil impacts gene architecture

To validate the previous observations we expressed *E.coli* DNA topoisomerase I (TopA). TopA expression in *top1top2* mutants depletes negative supercoil in plasmids[26]. wt and *top1Δ top2-1* cells harboring either control vector or TopA expressing plasmids were analyzed at 60 and 120 mins at the restrictive temperature for *top2-1* mutation (Fig. 3a,b). TopA expression in wt cells showed a reduction of negative supercoil at ORF flanking regions and, in *top1Δ top2-1* cells, nearly abolished the negative supercoil at flanking regions (Fig. 3a and Extended Data Fig. 5a). Hence, the presence of Hmo1 at gene boundaries in *top1Δ top2-1* double mutants does not prevent TopA from resolving negative supercoil. TopA, acts on negative supercoil converting it into positive supercoil[26]. Accordingly, the disappearance of negative supercoils at flanking regions paralleled the progressive accumulation of overwound DNA at the same location (Fig. 3b). TopA expression caused a reduction of positive supercoil at transcribed regions in wt cells, while it had the opposite effect in *top1Δ top2-1* mutants (Fig. 3b). This could result from the diffusion of supercoil waves across the entire gene bodies, perhaps due to the destruction of the topological/architectural confinements. Hmo1 binding was reduced in *top1Δ top2-1*compared to wt cells and nearly abolished in *top1Δ top2-1*expressing TopA (Extended Data Fig. 5b), indicating that Hmo1 association with gene boundaries depends on negative supercoil.

In wt cells, histone-H3 were distributed at transcribed units but was less abundant at gene boundaries (Fig 3c). *top1top2* mutants resembled wt cells, suggesting that the aberrant topological context in the double mutants did not affect the nucleosome context. TopA expression did not alter nucleosome positioning and distribution in wt cells, but in *top1top2* mutants, caused reduction of H3 distribution (Extended Data Fig. 5c). Moreover, H3 redistributed as its levels increased at flanking regions and, concomitantly, decreased at transcribed units, starting from position +2 (Fig. 3c and Extended Data Fig. 5d). Hence, TopA expression caused an increase of positive supercoil followed by diffusion of supercoil waves across the entire gene body and massive nucleosome repositioning.

Top2 mediate chromatin loop formation

Using the ChIA-PET (Chromatin interaction analysis by paired-end tag sequencing)[27] method, we investigated whether Top2 mediate the formation of chromatin loops. We used Top2 as a bait in S phase cells. Following DNA sequencing, 1 million independently mapped PETs (Extended Data Fig 6a,b) were acquired and, by keeping 1kb minimum distance, we obtained 1887 inter ligation PET clusters (Extended Data Fig 6b,c). The Top2-mediated loop length varied in size, some being even larger than 10 KB (~100 interactions), while the majority of loops were between 1500 to 2000 bp size with a median being 1900 bp (Fig 4a,b and Extended Data Fig 6d). 64% of the interactions corresponded to previously described Top2 binding sites[3] and, for the majority of the interactions (66%), Top2 was found only at one end of the loop (Extended Data Fig 6e). Overall, 45% of the Pol II genes were located within loops. Several loops were organized in clusters with 31% loops containing more than one genes and 51% loops containing a single gene (Fig 4b, and Extended Data Fig 6e,f).

Discussion

We showed that Top1 localizes at coding regions. Top2 instead acts at negatively supercoiled gene boundaries and engages genes in loop-like structures, placing promoters and terminators into proximity[28]. Multiple twin topological domains likely generate within the gene loops by waves of Pol II complexes[29]. While in prokaryotes negative supercoil behind the first transcription bubble might adsorb the positive supercoil generated by the next approaching Pol II complex[30], in eukaryotes, it may enable nucleosome assembly following Pol II passage[31]. Hence, eukaryotic RNA polymerase progression might strongly depend on Top1 in resolving topological stress in front of Pol II. Accordingly, Top1 accumulation at coding regions depends on transcription levels. Coding regions exhibit a positive supercoiled context, even when transcription is repressed, implying that genes retain a "memory" of a topological architecture that does not reflect the dynamics of elongating Pol II[7,32].

The negatively supercoiled regions flanking ORFs are refractory to nucleosome formation; TopA depleted negative supercoil specifically at these regions. Hence, gene boundaries exhibit an ideal topological context to breath out and undergo alternative structural transitions[33]. Nucleosome-free negatively supercoiled regions can form pseudo-cruciform structures[34], characterized by two B-DNA duplex arms and two intra-strand plectonemic arms in a non-B DNA conformation (Extended Data Fig. 7a). Such structures can branch migrate modulating the extension of the intra-strand plectonemic duplexes[35]. Like other HMG box proteins[36], Hmo1 binds four-ways junctions with high affinity. Moreover, it stabilizes nucleosome free regions, and dimerizes promoting DNA bridging[37]. We propose that Hmo1 locks cruciform DNA counteracting its branch migration and nucleosome formation. Stable negatively supercoiled gene boundaries in a cruciform conformation might contribute to insulate the topological architecture of gene loops to facilitate elongation of the multiple Pol II complexes, allowing efficient recycling of Pol II from TTS to TSS. Hmo1 dimerization activity[37] may promote gene looping, even in G1, and without the mediation of Top2. Pol II movement and transcription-coupled processes, such as gene gating and/or splicing, might also contribute to gene looping by extruding portions of the transcribed DNA[38]. In S phase, Top2 would act at the loop base, likely to counteract the disruptive potential of incoming forks, and/or to reset gene topology after fork passage (Extended Data Fig. 7b). Intriguingly, Top2-dependent DNA loops can contain more than one transcription unit, and can be organized in clusters, thus generating complex topological structures.

S phase cells accumulate RNA-DNA hybrids at ORFs in ~45% of the genes, independently of gene expression levels and of transcription-replication direction. However, the intergenic regions of converging genes exhibit a bias for hybrid accumulation. Pol II backtracking during elongation, and Pol II slow down during termination[39,40] could account for the formation of hybrids at ORFs under physiological conditions. The under-wound DNA behind Pol II can easily accommodate RNA:DNA pairing[41], and might even muffle the negative supercoil generated by Pol II movement. Our data

suggest that hybrid formation is a physiological event, intrinsic to the topological dynamics generated by transcription and co-transcriptional processes. However, converging genes might generate the context for unscheduled genotoxic events, as in the case of CSR-activated B cells[42].

Our model (Extended Data Fig. 7b) has the following predictions. i) Hmo1recruitment at gene boundaries would depend on their negative supercoil state; Hmo1 would then generate stable negative supercoiled cruciforms at gene boundaries (Extended Data Fig. 7a). Hmo1 is always found at negative supercoiled and nucleosome free regions flanking ORFs, while Top2 is recruited in S phase. Counteracting negative supercoil at gene boundaries prevents Hmo1 recruitment. Without Hmo1, cruciforms would be unstable but remain in a negative supercoiled state, becoming an ideal substrate for Top2 (Extended Data Fig. 7c). Accordingly, Top2 inactivation in *hmo1* cells rescues negative supercoil at gene boundaries. ii) in *top2* mutants negative supercoil decreases at ORF flanking regions, likely due to the unscheduled and massive recruitment of Top1 at gene boundaries. Notably, non-B DNA structures, can be a substrate for Top1[43] and Top1 can relax both positive and negative supercoil efficiently[44]. Hence, Hmo1 cannot protect cruciforms from Top1 activity when *TOP2* is mutated, implying that, in *top2* mutants, Top1 might cause genotoxic events at the Hmo1-locked cruciforms, such as extensive nicking and/or knotting[45] (Extended Data Fig. 7c). Intriguingly, *HMO1* deletion in *top2* mutants, besides alleviating *top2* temperature sensitivity[3], prevents Top1 relocation at flanking regions and *hmo1top2* double mutants exhibit a wt-like topological context. iii) In S phase *top2* mutants Pol II leaks outside the canonical transcribed regions. This aberrant Pol II distribution likely reflects the inability of *top2* mutants to recycle Pol II from TTS to TSS sites, due to the loss of proximity between promoters and terminators. In this view, Top2 might protect the gene loop structure from incoming forks. iv) The aberrant Pol II distribution in *top2* mutants may also account for hybrid accumulation at gene boundaries. In *top2* mutants, hybrid accumulation downstream of ORFs may result from aberrant transcription termination, while upstream of ORFs might be facilitated by the Top1-mediated processing of cruciform DNA. In fact, *top2 top1* exhibit less hybrids compared to *top2* alone. Another possibility is that Top2 defects promote aberrant antisense transcription initiation events close to TSSs.

The hybrids accumulating at gene boundaries in *top2* mutants may generate genotoxic events and the unscheduled synthesis of small RNA species, and might contribute to absorb the negative supercoil, thus implying that negative supercoil reduction at flanking regions may represent a non-direct consequence of Top1 relocation.

Our observations suggest that Top1, Top2 and Hmo1 contribute to the topological architecture of transcribed genes, particularly in S phase when forks reset the topological state of chromosomes and their chromatin context. Interfering with the topological context of gene flanking regions may cause a variety of pathological consequences such as the generation of aberrant RNA species, the accumulation of RNA:DNA hybrids and alterations at the level of chromatin architecture.

REFERENCE

1 Bermejo, R., Lai, M. S. & Foiani, M. Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Mol Cell* **45**, 710-718, doi:10.1016/j.molcel.2012.03.001 (2012).

2 Garcia-Muse, T. & Aguilera, A. Transcription-replication conflicts: how they occur and how they are resolved. *Nature reviews. Molecular cell biology* **17**, 553-563, doi:10.1038/nrm.2016.88 (2016).

3 Bermejo, R. *et al.* Genome-organizing factors Top2 and Hmo1 prevent chromosome fragility at sites of S phase transcription. *Cell* **138**, 870-884, doi:10.1016/j.cell.2009.06.022 (2009).

4 Sperling, A. S., Jeong, K. S., Kitada, T. & Grunstein, M. Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II

in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12693-12698, doi:10.1073/pnas.1106834108 (2011).

5 Wang, J. C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nature reviews. Molecular cell biology* **3**, 430-440, doi:10.1038/nrm831 (2002).

6 Pedersen, J. M. *et al.* DNA Topoisomerases Maintain Promoters in a State Competent for Transcriptional Activation in Saccharomyces cerevisiae. *PLoS genetics* **8**, doi:ARTN e100312810.1371/journal.pgen.1003128 (2012).

7 Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7024-7027 (1987).

8 Postow, L., Crisona, N. J., Peter, B. J., Hardy, C. D. & Cozzarelli, N. R. Topological challenges to DNA replication: conformations at the fork. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8219-8226, doi:10.1073/pnas.111006998 (2001).

9 Schvartzman, J. B. & Stasiak, A. A topological view of the replicon. *Embo Rep* **5**, 256-261, doi:10.1038/sj.embor.7400101 (2004).

10 Lal, A. *et al.* Genome scale patterns of supercoiling in a bacterial chromosome. *Nat Commun* **7**, 11055, doi:10.1038/ncomms11055 (2016).

11 Bermudez, I., Garcia-Martinez, J., Perez-Ortin, J. E. & Roca, J. A method for genome-wide analysis of DNA helical tension by means of psoralen-DNA photobinding. *Nucleic acids research* **38**, e182, doi:10.1093/nar/gkq687 (2010).

12 Naughton, C. *et al.* Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature structural & molecular biology* **20**, 387-395, doi:10.1038/nsmb.2509 (2013).

13 Kouzine, F. *et al.* Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nature structural & molecular biology* **20**, 396-403, doi:10.1038/nsmb.2517 (2013).

14 Sinden, R. R., Carlson, J. O. & Pettijohn, D. E. Torsional tension in the DNA double helix measured with trimethylpsoralen in living E. coli cells: analogous measurements in insect and human cells. *Cell* **21**, 773-783 (1980).

15 Perales, R. & Bentley, D. "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36**, 178-191, doi:10.1016/j.molcel.2009.09.018 (2009).

16 Bermejo, R. *et al.* Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes & development* **21**, 1921-1936, doi:10.1101/gad.432107 (2007).

17 Boguslawski, S. J. *et al.* Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods* **89**, 123-130 (1986).

18 Chan, Y. A. *et al.* Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS genetics* **10**, e1004288, doi:10.1371/journal.pgen.1004288 (2014).

19 Hamperl, S. & Cimprich, K. A. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA repair* **19**, 84-94, doi:10.1016/j.dnarep.2014.03.023 (2014).

20 Rossi, S. E., Ajazi, A., Carotenuto, W., Foiani, M. & Giannattasio, M. Rad53-Mediated Regulation of Rrm3 and Pif1 DNA Helicases Contributes to Prevention of Aberrant Fork Transitions under Replication Stress. *Cell Rep* **13**, 80-92, doi:10.1016/j.celrep.2015.08.073 (2015).

21 Hazelbaker, D. Z., Marquardt, S., Wlotzka, W. & Buratowski, S. Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol Cell* **49**, 55-66, doi:10.1016/j.molcel.2012.10.014 (2013).

22 Skourti-Stathaki, K., Kamieniarz-Gdula, K. & Proudfoot, N. J. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436-439, doi:10.1038/nature13787 (2014).

23 Tuduri, S. *et al.* Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature cell biology* **11**, 1315-1324, doi:10.1038/ncb1984 (2009).
24 Fachinetti, D. *et al.* Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. *Mol Cell* **39**, 595-605, doi:10.1016/j.molcel.2010.07.024 (2010).
25 Chen, S., Reger, R., Miller, C. & Hyman, L. E. Transcriptional terminators of RNA polymerase II are associated with yeast replication origins. *Nucleic acids research* **24**, 2885-2893 (1996).
26 Gartenberg, M. R. & Wang, J. C. Positive Supercoiling of DNA Greatly Diminishes Messenger-Rna Synthesis in Yeast. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11461-11465, doi:DOI 10.1073/pnas.89.23.11461 (1992).
27 Li, X. *et al.* Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions. *Nat Protoc* **12**, 899-915, doi:10.1038/nprot.2017.012 (2017).
28 O'Sullivan, J. M. *et al.* Gene loops juxtapose promoters and terminators in yeast. *Nature genetics* **36**, 1014-1018, doi:10.1038/ng1411 (2004).
29 Levens, D., Baranello, L. & Kouzine, F. Controlling gene expression by DNA mechanics: emerging insights and challenges. *Biophysical reviews* **8**, 259-268, doi:10.1007/s12551-016-0216-8 (2016).
30 Rovinskiy, N., Agbleke, A. A., Chesnokova, O., Pang, Z. & Higgins, N. P. Rates of gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. *PLoS genetics* **8**, e1002845, doi:10.1371/journal.pgen.1002845 (2012).
31 Patterton, H. G. & von Holt, C. Negative supercoiling and nucleosome cores. I. The effect of negative supercoiling on the efficiency of nucleosome core formation in vitro. *Journal of molecular biology* **229**, 623-636, doi:10.1006/jmbi.1993.1068 (1993).
32 Ma, J., Bai, L. & Wang, M. D. Transcription under torsion. *Science* **340**, 1580-1583, doi:10.1126/science.1235441 (2013).
33 Kouzine, F. *et al.* Permanganate/S1 Nuclease Footprinting Reveals Non-B DNA Structures with Regulatory Potential across a Mammalian Genome. *Cell Syst* **4**, 344-356 e347, doi:10.1016/j.cels.2017.01.013 (2017).
34 Lilley, D. M. DNA opens up--supercoiling and heavy breathing. *Trends Genet* **4**, 111-114 (1988).
35 Murchie, A. I. & Lilley, D. M. Supercoiled DNA and cruciform structures. *Methods in enzymology* **211**, 158-180 (1992).
36 JR, P. o., Norman, D. G., Bramham, J., Bianchi, M. E. & Lilley, D. M. HMG box proteins bind to four-way DNA junctions in their open conformation. *Embo J* **17**, 817-826, doi:10.1093/emboj/17.3.817 (1998).
37 Murugesapillai, D. *et al.* DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin. *Nucleic acids research* **42**, 8996-9004, doi:10.1093/nar/gku635 (2014).
38 Tan-Wong, S. M., Wijayatilake, H. D. & Proudfoot, N. J. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes & development* **23**, 2610-2624, doi:10.1101/gad.1823209 (2009).
39 Saponaro, M. *et al.* RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* **157**, 1037-1049, doi:10.1016/j.cell.2014.03.048 (2014).
40 Cheung, A. C. & Cramer, P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249-253, doi:10.1038/nature09785 (2011).
41 Drolet, M. *et al.* Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of

DNA topoisomerase I. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3526-3530 (1995).
42 Meng, F. L. *et al.* Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* **159**, 1538-1548, doi:10.1016/j.cell.2014.11.014 (2014).
43 Husain, A. *et al.* Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. *Nat Commun* **7**, 10549, doi:10.1038/ncomms10549 (2016).
44 Fernandez, X., Diaz-Ingelmo, O., Martinez-Garcia, B. & Roca, J. Chromatin regulates DNA torsional energy via topoisomerase II-mediated relaxation of positive supercoils. *Embo J* **33**, 1492-1501, doi:10.15252/embj.201488091 (2014).
45 Brown, P. O. & Cozzarelli, N. R. Catenation and knotting of duplex DNA by type 1 topoisomerases: a mechanistic parallel with type 2 topoisomerases. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 843-847 (1981).

FIGURE legends

Fig 1. Topological context of Pol II genes. ChIP-on-chip was carried out in cells released from G1 into S. Pol II coding regions (replicates n=2; meta-gene analysis n=6706 genes) were scaled to 1kbp and flanking 0.5kbp from TSS and TTS were plotted against median intensity on the Y-axis. (a) Meta-gene plot showing accumulation of Pol II (Rpb3-10X Flag), Top2 (Top2-10X Flag), Top1 (Top1-10X Flag) and Hmo1(Hmo1-10X Flag). (b) Meta-gene plot for negative supercoil in G1 and S, plotted against median bTMP intensity. (c) Positive supercoil distribution in Pol II genes were plotted against average gene density on the Y-axis. (d) Pol II genes were grouped into three categories; -high, -medium and –low expression genes based on FPKM (Fragments Per Kilobase Million) value from RNA-seq carried out in S-phase at 28^0C. Meta-gene plots for three categories of gene expression for Pol II, Top2, Top1 and Hmo1. (e) Negative supercoil distribution in high, medium and low expression genes. (f) Pol II genes were grouped according to their orientation with respect to neighbouring genes as: Co-directional (+ strand; n=1453 gene pairs), Co-directional (- strand; n=1415 gene pairs), Converging (n=1590 gene pairs) and Diverging (n=1512 gene pairs). Number of gene pairs at different intergenic space (<250 bp = 1729 gene pairs, 251-500bp = 2224 gene pairs and >500 bp = 2010 gene pairs) were plotted with respect to their orientation. (g) Base coverage percentage of supercoil accumulation at different intergenic space with respect to gene pairs grouped according to orientation (2 replicates, mean ± s.d).

Fig 2. Top2 and Hmo1 contributes to gene topology. G1 cells were released at 28^0C into S and temperature shifted to 37^0C for *top2-1* mutants (replicates n=2; meta-gene analysis n=6706 genes). (a) Meta-gene profiles for negative supercoil comparison in wild type, *top2-1*, *top1*Δ and *top2-1top1*Δ cells. (b) Meta-gene profiles comparing wt and *top2-1* cells for Top1 protein (Top1-10X Flag) accumulation and (c) Pol II (Rpb3-10X Flag) accumulation. (d) Meta-gene profile comparison for negative supercoil in wt, *hmo1*Δ and *top2-1hmo1*Δ cells. (e) Meta-gene profiles for RNA:DNA hybrid comparison in wt, *top2-1*and *top1*Δ cells. (f) Percentage of genes (G) and RNA:DNA hybrid (H) in either head-on or co-directional with respect to replication forks. Genes were grouped with respect to their distance (0.25kbp, n=140 genes; 0.5kbp, n=235 genes; 1kbp, n=347 genes; 2kbp, n=539 genes and 5kbp, n=1121 genes) and direction (head-on or co-directional) from replication origin.

F**ig 3. Negative supercoil disruption causes disarray in nucleosome occupancy.** Wt and *top1*Δ *top2-1* mutants harbouring either control plasmid or expressing *E.coli* DNA topoisomerase I (TopA) plasmid were grown at 28^0C and shifted to 37^0C for 60 or 120 mins to inactivate Top2 (replicates n=2; meta-gene analysis n=6706 genes). (a) Meta-gene profiles for negative supercoil comparison in wt [Control], wt[TopA], *top2-1top1*Δ [Control] and *top2-1top1*Δ [TopA] at 60 mins and 120 mins at

restrictive temperature. (b) Meta-gene profile for positive supercoil accumulation. (c) Meta-gene profiles of histone H3 in wt and *top2-1 top1*Δ mutants with [Control] or [TopA] plasmids were plotted against median read coverage on the Y-axis.

Fig 4. Top2 mediates chromatin loop formation.
(a) Box plot showing loop size distribution (n=1505 loops; min=1.1kb; max=662kb; median=1.9kb; 25th Percentile=1.6kb; 75th Percentile=2.4kb) (b) Genome browser view of Top2 mediates chromatin interactions on chromosome II, along with Top2 and Hmo1 protein chip data. Highlighted area (Chr II: 242000 to 258000) is enlarged in the lower panel.

METHODS

Strains and growing conditions

All *S. cerevisiae* strains are W303 derivatives[46]. The relevant genotypes are shown in Extended Data Table 1. Strains were grown at 28 °C YPD medium. G1 synchronization was carried out using 3-5 μg/ml of α factor. For S-phase samples, G1 cells were washed twice in YP medium and allowed to grow for 15 min in fresh medium. For temperature-sensitive strains, cells were allowed to grow for 10 min in fresh medium after G1 release, centrifuged and then dissolved in pre-warmed media at 37 °C and allowed to grow for 15 min. Cell cycle progression into S phase was monitored by FACS and budding profiles. For *E. coli* TopA expression, wild type and *top1Δ top2-1* mutants harbouring either control or TopA expression plasmids were grown at 25 °C in synthetic medium lacking leucine. Cells were shifted to 37 °C for inactivation of Top2 after reaching 8×10^6 cells/ml concentration.

bTMP ChIP

We adopted the previously described method to yeast[12]. Sodium azide (0.1%) was used to block cells and to ensure the preservation of the most prevalent topological context present at any given genomic position. We note that this method does not aim at studying dynamic topological transitions. Permeabilized yeast cells were incubated with bTMP (400 μg per $2x10^9$ cells) in dark for 90 min and then cross-linked by 365 nm UV (800 mJ/cm^2) light to form adducts between two DNA strands. Cells were washed twice with ice-cold PBS and lysed in 1 ml of lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) using Zirconia beads. The crosslinked chromatin was sheared to an average size of 500 bp by 6 x 15-sec pulses using a Biorupter sonicator and DNA was purified. Purified DNA was incubated with Dynabeads MyOne streptavidin (Invitrogen cat. no. 65001) overnight at 4 °C . The beads were washed 2x with each of the following buffers; wash buffer-I (20 mM Tris-HCL pH 8, 2 mM EDTA, 150 mM NaCl, 1% triton X, 0.1% SDS), wash buffer-II (20 mM Tris-HCL pH 8, 2 mM EDTA, 500 mM NaCl, 1% triton X, 0.1% SDS) wash buffer III (250 mM LiCl, 10 mM Tris pH 8.0, 0.5% Na-deoxycholate, 0.5% NP-40, 1 mM EDTA) and 1x TE (20 mM Tris pH 8.0, 2 mM EDTA). The bTMP-DNA complexes were eluted from the beads in 250 μl elution buffer (95% formamide, 10 mM EDTA) at 90 °C for 20 min and eluted samples were cleaned with Qiagen PCR clean up kit. Input DNA was isolated from sheared chromatin input (1/100 of the material used for ChIP). For bTMP-chip with naked DNA, genomic DNA was isolated from Qiagen Genomic-tip 100/G (cat. no. 13343) & Genomic DNA buffer Set (cat. no. 19060). Purified DNA was sheared to an average size of 500 bp by 6x 15-sec pulses using a Biorupter sonicator. bTMP was added to purified DNA and incubated in dark for 90 min and crosslinked with UV at 365 nm (800 mJ/cm^2). DNA was precipitated using isopropanol and washed with 70% ethanol. Dried pellet was dissolved in buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.1% SDS) and incubated with Dynabeads MyOne streptavidin (Invitrogen cat. no. 65001) overnight at 4 °C . Washing and elution was the same mentioned above. Both IP and Input samples were processed as mentioned in microarray section.

The procedures for bTMP titration is presented in Extended Data Figures 8. bTMP binding normalization and the dispersion profile for bTMP are presented in Extended Data Figure 9.

Protein ChIP

ChIP analysis for proteins was carried out as described[47] with few modifications. Cells were crosslinked with 1% formaldehyde in culture medium for 30 min at room temperature followed by quenching with 0.125 M glycine for 5 min. Cells were washed twice with ice-cold PBS and lysed in 1 ml of lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) using Zirconia beads. Crosslinked chromatin was sheared to an average size of 500 bp by 6x 15-sec pulses using a Biorupter sonicator. The lysate was then centrifuged to remove cell debris. The chromatin fraction was incubated with Dynabeads protein G beads (Invitrogen, cat. no. 10003D) coated with anti-Flag antibody (M2-antiflag, Sigma) overnight at 4 °C. The immune complexes were washed with the following buffers 2X; Chip-lysis buffer, high-salt lysis buffer (Chip-lysis buffer + 360 mM NaCl), Chip-wash buffer (250 mM LiCl, 10 mM Tris pH 8.0, 0.5% Na-deoxycholate, 0.5% NP-40, 1 mM EDTA) and 1x TE (20 mM Tris pH 8.0, 2 mM EDTA). The protein-DNA complexes were eluted from the beads in 250 µl elution buffer (1% SDS, 50 mM Tris pH 8.0, 10 mM EDTA) at 65 °C for 20 min followed by the addition of proteinase K to 500 µg/ml and overnight incubation at 65 °C . Input DNA was isolated from sheared chromatin input (1/100 of the material used for ChIP). Both IP and Input samples were processed as mentioned in microarray section.

DRIP-ChIP

DRIP-ChIP was performed using anti-DNA:RNA hybrid monoclonal mouse antibody S9.6 as previously described[18]. Briefly, cells were crosslinked with 1% formaldehyde in culture medium for 20 min at room temperature followed by quenching with 0.125 M glycine for 5 min. Cells were washed twice with ice-cold PBS and lysed in 1 ml of lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) using Zirconia beads. Crosslinked chromatin was sheared to an average size of 500 bp by 6x 15-sec pulses using a Biorupter sonicator. The lysate was then centrifuged to remove cell debris. Chromatin fraction was incubated with Protein-A magnetic beads (Invitrogen, cat. no. 10001D) coated with anti-DNA:RNA hybrid S9.6 antibody[17] overnight at 4 °C . The immune complexes were washed with the following buffers 2X; Chip-lysis buffer, high-salt lysis buffer (Chip-lysis buffer + 360 mM NaCl), Chip-wash buffer (250 mM LiCl, 10 mM Tris pH 8.0, 0.5% Na-deoxycholate, 0.5% NP-40, 1 mM EDTA) and 1x TE (20 mM Tris pH 8.0, 2 mM EDTA). The RNA:DNA hybrid complexes were eluted from the beads in 250 µl elution buffer (1% SDS, 50 mM Tris pH 8.0, 10 mM EDTA) at 65 °C for 20 min followed by the addition of proteinase K to 500 µg/ml and overnight incubation at 65 °C. Input DNA was isolated from sheared chromatin input (1/100 of the material used for ChIP). Both IP and Input samples were processed as mentioned in microarray section.

Histone H3 ChIP-seq

ChIP analysis for proteins was carried out as described previously[48]. Briefly, cells were crosslinked with 1% formaldehyde in culture medium for 15 min at room temperature followed by quenching with 0.125 M glycine for 5 min. Cells were washed twice with ice-cold PBS and lysed in 1 ml of lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) using Zirconia beads. Crosslinked chromatin was sheared to an average size of 200 bp in Covaris S220 Focused-ultrasonicators. The lysate was then centrifuged to remove cell debris. Chromatin fraction was incubated with Protein-G magnetic beads (Invitrogen, cat. no. 10003D) coated with Anti-Histone H3 antibody (Abcam, cat. no. ab1791) overnight at 4 °C . The immune complexes were washed with the following buffers 2X; Chip-lysis buffer, high-salt lysis buffer (Chip-lysis buffer + 360 mM NaCl), Chip-wash buffer (250 mM LiCl, 10 mM Tris pH 8.0, 0.5% Na-

deoxycholate, 0.5% NP-40, 1 mM EDTA) and 1x TE (20 mM Tris pH 8.0, 2 mM EDTA). The protein-DNA complexes were eluted from the beads in 250 µl elution buffer (1% SDS, 50 mM Tris pH 8.0, 10 mM EDTA) at 65 °C for 20 min followed by the addition of proteinase K to 500 µg/ml and overnight incubation at 65 °C. Input DNA was isolated from sheared chromatin input (1/100 of the material used for ChIP).
For sequencing, IP and Input ChIP-seq libraries were prepared according to the manufacturer protocols for the Ion Proton sequencer (Thermo Fisher Scientific/Life Technologies). Briefly 10 ng for ChIP DNA was end repaired and adapter ligated using the KAPA Library Preparation Kit for Ion Torrent™ (KAPABIOSYSTEMS, inc) and adapter barcode Kapa Barcode Adaptors 9-24. After adapter ligation, each sample was size selected using AMPure XP Bead (Beckman Coulter, inc). An amplification reaction was set up in a final volume of 50 ul. A SPRI cleanup with a 1.5X Bead:DNA ratio was performed post amplification and final libraries were eluted in 35 ul. Libraries were quantified on Qubit fluorometer with HS DNA (Thermo Fisher Scientific/Life Technologies) and checked for size on an Agilent Bioanalyzer with HS DNA kit (Agilent, Santa Clara, CA). Each size selected library was diluted according to the final concentration of 11 pM and clonally amplified using the Ion Proton™ Hi-Q™ Template Kit (Thermo Fisher Scientific/Life Technologies) with IonOneTouch 2 instrument (Thermo Fisher Scientific/Life Technologies). After emulsion PCR, DNA positive ISPs were recovered and enriched according to standard protocols with the IonOneTouch ES Instrument (Thermo Fisher Scientific/Life Technologies). A sequencing primer was annealed to DNA positive ISPs and the sequencing polymerase bound, prior to loading of ISPs into Ion P1 sequencing chips. Sequencing of the samples was conducted according to the Ion Proton™ Hi-Q Sequencing Kit Protocol. One P1 sequencing chips with 6 libraries were loaded and run on an Ion Proton sequencer.

RNA-seq
Total RNA was isolated from $5x10^7$ cells by using RNeasy mini kit (50) (Qiagen cat. no. 74104). Prior to library preparation cytoplasmic and mitochondrial ribosomal RNA was removed using 'Ribo-Zero Gold rRNA Removal Kit (Yeast)' (Illumina, cat. no. MRZY1324). Libraries for RNA sequencing were prepared following the manufacturer protocols for transcriptome sequencing with the Ion Proton sequencer (Thermo Fisher Scientific/Life Technologies). Briefly, 1 µg of total RNA was poly-A selected using the Dynabeads mRNA Direct Micro Purification kit (ThermoFisher, cat. no. 61021) according to manufacturer's protocol. About 50 ng of poly-A RNA were used to prepare strand-specific barcoded RNA libraries with the Ion Total RNA-Seq kit v2.0 (ThermoFisher Scientific, cat. no. 4475936). Briefly, poly-A RNA was fragmented with RNase III and purified with Nucleic Acid Binding Beads. After purification, the poly-A RNA fragmented were hybridized and ligated with Ion Adaptor and subsequently reverse transcribed for cDNA preparation. cDNAs were amplified with Ion Torrent barcoded primers and purified with Nucleic Acid Binding Beads. Final libraries were quantified on Qubit fluorometer with HS DNA (Thermo Fisher Scientific) and checked for size on an Agilent Bioanalyzer with HS DNA kit (Agilent, Santa Clara, CA). Four barcoded libraries were pooled together on an equimolar basis according to the final concentration of 11 pM and clonally amplified using the Ion Proton™ Hi-Q™Template Kit (Thermo Fisher Scientific, cat no. A26434) with IonOneTouch 2 instrument (Thermo Fisher Scientific/Life Technologies). After emulsion PCR, DNA positive ISPs were recovered and enriched according to standard protocols with the IonOneTouch ES Instrument (Thermo Fisher Scientific/Life Technologies). A sequencing primer was annealed to DNA positive ISPs and the sequencing polymerase bound, prior to loading of ISPs into Ion P1 sequencing chips. Sequencing of the samples was conducted according to the Ion Proton™ Hi-Q Sequencing Kit (Thermo Fisher Scientific, cat. no. A26433) Protocol on Ion Proton instrument.

ChIA-PET (Chromatin interaction analysis by paired-end tag sequencing)
We adopted the previously described method[27]. Cells were crosslinked with 1% formaldehyde in culture medium for 30 min at room temperature followed by quenching with 0.125 M glycine for 5

min. Cells were washed twice with ice-cold PBS and lysed in 1 ml of lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) using Zirconia beads. Crosslinked chromatin was sheared to an average size of 500 bp by 6x 15-sec pulses using a Biorupter sonicator. The lysate was then centrifuged to remove cell debris. The chromatin fraction was incubated with Dynabeads protein G beads (Invitrogen, cat no 10003D) coated with anti-Flag antibody (M2-antiflag, Sigma) overnight at 4 °C . The immune complexes were washed with the following buffers 2X; Chip-lysis buffer, high-salt lysis buffer (Chip-lysis buffer + 360 mM NaCl), Chip-wash buffer (250 mM LiCl, 10 mM Tris pH 8.0, 0.5% Na-deoxycholate, 0.5% NP-40, 1 mM EDTA) and 1x TE (20 mM Tris pH8.0, 2 mM EDTA). Beads were pooled and end-repair was carried out using T4 DNA polymerase (NEB, cat. no. M0203L) by rotating on a Intelli-Mixer at 37 °C for 40 min. Beads were washed 3x with ice cold ChIA-PET wash buffer (10 mM Tris pH 7.4, 1 mM EDTA, 500 mM NaCl). A-tailing was carried out using Klenow fragment (3'-5' exo-) (NEB, cat. no. M0212M) in the presence of 100 uM dATP by rotating on a Intelli-Mixer at 37 °C for 50 min. Beads were washed 3x with ice cold ChIA-PET wash buffer (10 mM Tris pH 7.4, 1 mM EDTA, 500 mM NaCl). For proximity ligation, a bridge linker was prepared by annealing Linker-F and Linker-R (HPLC purified (250nmole) from IDT (Integrated DNA Technologies)):

Bridge linker-F: 5'- /5Phos/CGCGATATC/iBIOdT/TATCTGACT -3'
Bridge linker-R: 5'- /5Phos/GTCAGATAAGATATCGCGT -3'.

Proximity ligation was carried out by T4 DNA ligase (NEB cat. no. M0202M), in the presence of bridge linker at a concentration of 0.57 ng/ul by rotating at 16 °C overnight. Beads were washed once in ChIA-PET wash buffer and eluted in elution buffer (10 mM Tris pH 8.0, 1 mM EDTA, 1% SDS) at 65 °C for 15 min. Reverse crosslinking was carried out at 65 °C in the presence of proteinase K. DNA was purified using phenol-chloroform-isoamyl alcohol (pH 7.9) and MaXtract High Density-2ml (QIAGEN, cat. no. 129056) and precipitated with isopropanol. Tagmentation of proximity ligated DNA was carried by Tn*5* transposome by using Nextera® DNA Sample Preparation Kit (24) (Illumina FC-121-1030). Tagmentated DNA was purified using Zymo Genomic DNA Clean & Concentrator™ kit (Zymo Research, cat. no. D4014) and fragments containing linker DNA was enriched with Dynabeads M-280 Streptavidin (Invitrogen, cat. no.11205D). Beads were washed with 2×SSC/0.5% (wt/vol) SDS five times and twice in 1X B&W buffer (10 mM Tris pH 8.0, 1 mM EDTA, 1 M NaCl). Sequencing library was amplified by using beads and purified library was used for pair-end sequencing using MiSeq® Reagent Kit v3 (600 cycle) on Illumina MiSeq instrument.

Microarray and data processing

Both IP and Input DNA was amplified using GenomePlex complete whole genome amplification kit (Sigma, cat. no. WGA1-50RXN) biotin labelled and hybridized to Affymetrix GeneChip *S. cerevisiae* Tiling 1.0R Array (Sc03b_MR) according to the Affymetrix standard protocol. The CEL files were processed using rMAT[49] R package to identify enriched regions across the genome. At first, systematic biases such as probe effect were corrected by normalization. Then probe intensities were smoothed and a score was calculated for each probe using IP and Input. To detect enriched regions based on the probe score the following parameters were used; dMax=300 (sliding window side), nProbesMin=8 (Minimum number of probes to average), method=Score (calling enriched regions based on sliding widow scores), $\log_2$ threshold=1.5 (equal and greater than 1.5 are labelled as enriched region). For bTMP experiments, bTMP binding 'in cells' (IP/Input) were subtracted by 'naked genomic DNA' score (IP/Input) to correct for false positives binding of bTMP.

Meta-gene analysis

Meta-gene was used to study the averaged enriched peak profile across all protein coding genes (6706 genes from sacCer 2011annotation) or specific set of genes with upstream (-500 bases from TSS) and downstream (+500 bases from TTS) in the yeast genome. The peak scores were mapped using

bedtools[50], for every base of the gene including upstream (-500b) and downstream (+500b). The length of the gene was scaled to 1000 bases. For scaling ORF region to 1000 bases, the following equation was iterated for every base across all the genes $((Z - x_i)/(y_i - x_i)) * 1000$; x_i = Start position of the i^{th} gene i = (1,2,3 … total genes), y_i = End position of the i^{th} gene i = (1,2,3 … total genes), Z = base position (1,2,3….$y_i - x_i$). For average intensity plot, the IP/Input values of the normalized position (1000 ORF, -500 Upstream and +500 Downstream) of each gene was aggregated using median. For average gene density, the IP/Input score was converted into a categorical value either 1 or 0 based on the threshold of 1.5 (>=1.5 is 1 and <1.5 is 0) of all the normalized position (1000 ORF, -500 upstream and +500 downstream) of each gene and aggregated with sum function. For visualization, average intensity and average gene density was plotted respective to normalized ORF position. The points were smoothed using generalized additive model (GAM) to obtain a curve using ggplot2 R package.

RNA-seq data processing

The RNA-seq data from IonTorrent proton instrument contains approximately 25 million reads for each sample. The raw reads were filtered based on the quality value (-q 20 and –p 30) using FASTX Toolkit. The filtered reads were aligned to the reference genome (SacCer 2011) using STAR aligner[51]. Aligned BAM files were used for transcript quantification (FPKM-Fragments Per Kilo Million) using RSEM[52]. The genes sets were divided into three equal categories (low, medium and high expression) based on the FPKM values and used to plot the supercoiling, protein and RNA:DNA hybrid profile using the meta-gene calculation mentioned above.

Histone H3 ChIP-seq data processing

The ChIP-seq data from IonTorrent proton instrument contains approximately 15 million reads for each sample. The raw reads were filtered based on the quality value (-q 20 and –p 30) using FASTX Toolkit. The filtered reads were aligned to the reference genome (SacCer 2011) using TMAP aligner. The PCR duplicates were removed from the aligned BAM files using PICARD tools. The BAM files were sorted and indexed for the peak calling using SAMTOOLS. The bedgraph files were generated by comparing bam files of IP and Input (IP read coverage/Input read coverage) result in the ratio for every base across the whole genome using DEEPTOOLS (bamCompare)[53]. Finally, peak calling was performed using DANPOS (dpos) toolkit[54] with the IP/Input threshold 1.4 (-q 1.4) where the output peaks corresponds to the individual nucleosome. The DANPOS was preferred over MACS toolkit for the dynamic nucleosome analysis at single nucleotide resolution.

ChIA-PET data processing

ChIA-PET data contains approximately 10 million reads with median length of approximately 105 nucleotides. Raw reads were filtered based on the quality value (-q 20 and –p 30) using FASTX Toolkit. The filtered reads were scanned for bridge linker (ACGCGATATCTTATCTGACT, AGTCAGATAAGATATCGCGT) with maximum two mismatches using cutadapt. The reads containing the bridge linker were aligned to the reference genome (SacCer 2011) using bwa mem module. PCR duplicates were removed using Picard Markduplicates module. The aligned bam file was converted to bed pair end interaction file (bedpe) for cluster generation using bedtools (bamtobed) module. Paired end tags (PET) with less than 1 kb distance (self-ligation loops) were not considered for the PET clustering. Individual PET interactions were clustered by extending each PET by 500 bp and PETs overlapping at both ends were clustered together as a single PET cluster[27]. PET clusters with more than or equal to 2 were considered for meta-analysis. WashU Epigenome Browser was used to visualize chromatin-chromatin interactions[55].

Tool kits

FASTX Toolkit: http://hannonlab.cshl.edu/fastx_toolkit/
TMAP Toolkit: https://github.com/iontorrent/TMAP

PICARD Toolkit: https://broadinstitute.github.io/picard/
BWA Toolkit: http://bio-bwa.sourceforge.net/

STATISTICS AND REPRODUCIBILITY

All the experiments were carried out with two biological replicates. To test the significance of the overlap between two replicates (supercoiling, protein and hybrid peak calls), intersect and Fisher exact test from bedtools were used. For bedtools intersect, minimum of 80% overlap were expected for further downstream analysis like meta-gene plotting. The number of overlap peaks and sum of overlap bases between two sets of intervals from bedtools were visualized using VennDiagram library from R. Protein coding genes (n=6706) from saccer2011 were used for meta-gene plotting.

ADDITIONAL REFERENCES

46 Thomas, B. J. & Rothstein, R. Elevated recombination rates in transcriptionally active DNA. *Cell* **56**, 619-630 (1989).

47 Bermejo, R., Katou, Y. M., Shirahige, K. & Foiani, M. ChIP-on-chip analysis of DNA topoisomerases. *Methods in molecular biology* **582**, 103-118, doi:10.1007/978-1-60761-340-4_9 (2009).

48 Rodriguez, J., McKnight, J. N. & Tsukiyama, T. Genome-Wide Analysis of Nucleosome Positions, Occupancy, and Accessibility in Yeast: Nucleosome Mapping, High-Resolution Histone ChIP, and NCAM. *Curr Protoc Mol Biol* **108**, 21 28 21-16, doi:10.1002/0471142727.mb2128s108 (2014).

49 Droit, A., Cheung, C. & Gottardo, R. rMAT--an R/Bioconductor package for analyzing ChIP-chip experiments. *Bioinformatics* **26**, 678-679, doi:10.1093/bioinformatics/btq023 (2010).

50 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

51 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

52 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323 (2011).

53 Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic acids research* **42**, W187-191, doi:10.1093/nar/gku365 (2014).

54 Chen, K. *et al.* DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome research* **23**, 341-351, doi:10.1101/gr.142067.112 (2013).

55 Zhou, X. *et al.* The Human Epigenome Browser at Washington University. *Nat Methods* **8**, 989-990, doi:10.1038/nmeth.1772 (2011).

ACKNOWLEDGEMNTS

We thank J. Roca (SBU-Barcelona) for sharing TopA expressing plasmids, and M.Bianchi (HSR-Milan), G.Liberi (CNR-PV) and all our lab members for discussions. We thank Cogentech and C. Valli, M. Riboni and S. Minardi for microarray and DNA sequencing. Research was supported by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC), the European Union, MIUR, Worldwide Cancer Research, and Telethon-Italy to MF. YJA is supported by European Community's Seventh Framewok Programme under Grant agreement no. 246549 – Train 2009. N.G. is funded by the UK Medical Research Council (MR/J00913X/1; MC_UU_00007/13).

AUTHOR CONTRIBUTIONS

Y.J.A. and M.F. designed the experiments, interpreted results, and prepared the manuscript. Y.J.A. and M.A. performed the experiments. M.A. performed statistical and computational analysis. R.C. provided technical inputs and N.G. provided bTMP and technical inputs for supercoil analysis.

COMPETING INTERSTS

The authors declare no competing interests.

CORRESPONDENCE AND REQUESTS FOR MATERIALS

should be addressed to YJA or M.F.

DATA AVAILABILITY

All the raw and processed data are available at the Gene Expression Omnibus (GEO) under following accession numbers:
GSE114410: bTMP, RNA:DNA hybrids, Top1 protein ChIP-chip, and RPB3 protein ChIP-chip
GSE114444: RNA-seq, H3 ChIP-seq and ChIA-PET
GSE16258 (Bermejo et al., 2009): Top2 protein ChIP-chip, Hmo1 protein ChIP-chip and RPB3 protein ChIP-chip

CODE AVAILABILITY

All the custom-made scripts used for this study are available in the GitHub repository at https://github.com/adhilmd/TopologyCustomAnalysis.

EXTENDED DATA LEGENDS

Extended Data Figure 1. DNA supercoil accumulation across the genome in wild-type conditions

(a) Schematic representation of bTMP binding to DNA in a supercoil dependent manner. Based on normalized bTMP (IP/Input) values, genomic region is categorized as Negative, Poitive and stable regions. (b) Genome browser profile for bTMP binding on chromosome-III. Based on peak intensities bTMP peaks were called for negative (red), stable (blue) and positive (green) regions. Side panel shows a zoomed region of chromome III spanning from 15 to 25 kb. For bTMP profile, positive value peaks (normalized IP/Input) which are above threshold (+1.5) were designated as 'Negatve supercoil' (-σ), and negative value peaks (normalized IP/Input) which were below threshold (-1.5) were designated as 'Positive supercoil' (+σ). Peaks in between threshold (from -1.5 to +1.5) were considered as stable regions. (c) Pie charts showing the coverage of negative, positive and stable regions based on bTMP-ChIP values were plotted as percentage coverage for whole genome, intergenic regions, protein coding regions or nucleosome-occupied regions. (d) Gene correlation plot for negative supercoil (percentage) accumulation for wild type cells in G1 and S-phase (n=6706 genes; paired t-test two-sided <2e-16; Pearson correlation r=0.86). (e) Venn diagram comparison of G1 and S-phase for bTMP binding with respect to peak number and base coverage. (f) Genome browser view for the *ASF2* locus on chromosome IV, showing bTMP peaks and Pol II (Rpb3-ChIP) accumulation in G1 and S phase. (g) Gal genes in Chromosome-II (spanning from 270 to 285 kb), depicting bTMP and Pol II (Rpb3-ChIP) binding profiles in Glucose and Galactose conditions in S-phase. (h) A zoomed view of the chromosome-III region spanning from 90 to 96 kb, containing a highly active gene (*LEU2*) close to tRNA (tL(CAA)C) and two moderately expressed *NFS1* and *DCC1* genes. Positive and negative supercoil and Pol II (Rpb3-ChIP), Top1, Top2 and Hmo1 accumulation are shown.

Extended Data Figure 2. Negative and positive supercoil distribution in DNA topoisomerase mutants

(a) Pol II coding genes were grouped into three categories; -high, -medium and –low expression genes based on FPKM (Fragments Per Kilobase Million) value from RNA-seq data (n=6706 genes; Low=Medium=High=2235 genes). (b) Positive supercoil distribution in high, medium and low expression genes. (c) Base coverage percentage of Top2-protein and Top1-protein accumulation at different intergenic space (<250 bp = 1729 gene pairs, 251-500bp = 2224 gene pairs and >500 bp = 2010 gene pairs) with respect to gene pairs grouped according to orientation. (d) Positive supercoil accumulation in wild type and *top2-1* cells in S-phase. Pol II genes were plotted against average gene density on the Y-axis. (e) Meta-gene plot for negative supercoil in in G1 synchronized cells in wild type and *top2-1*. (f) Positive supercoil accumulation in wild type and *top1*Δ cells in S-phase. (g) Positive supercoil accumulation in wild type and *top1*Δ*top2-1* cells in S-phase. (h) Genome browser profile of chromosome-III from region 90 to 96 kb, showing comparative bTMP binding in wild type and topoisomerase mutants. (i) Genome browser profile of chromosome-III from region 90 to 96 kb, showing Pol II (Rpb3-ChIP) and Top1 protein accumulation in wild type and *top2-1* mutants.

Extended Data Figure 3. Negative and positive supercoil accumulation in hmo1 mutants
(a) Positive supercoil accumulation in wild type and *hmo1*Δ cells in S phase. Pol II genes were plotted against average gene density on the Y-axis. (b) Positive supercoil accumulation in wild type and *top2-1hmo1*Δ cells in S-phase. (c) Top1 protein (Top1-10X Flag) accumulation comparison in wt, *hmo1*Δ and *top2-1hmo1*Δ cells. (d) Genome browser profile of chromosome-III from region 90 to 96 kb, showing comparative bTMP binding in wild type, *hmo1*Δ, and *top2-1hmo1*Δ cells in S-phase. (e) Genome browser profile of chromosome-III from region 90 to 96 kb, showing Top1 protein accumulation in wild type, *hmo1*Δ*, top2-1*, and *top2-1hmo1*Δ mutants.

Extended Data Figure 4. RNA:DNA hybrid accumulation in wild type and mutants
(a) Meta-gene profiles for RNA:DNA hybrid comparison in wild type (28^0C and 37^0C) (b) RNA:DNA hybrids accumulation in different expression classes. Pol II genes were grouped into three categories; -high, -medium and -low expression genes based on FPKM (Fragments Per Kilobase Million) value from RNA-seq (Extended Data Figure 1h). (c) Meta-gene profile for RNA:DNA hybrid in S-phase in wild type, *rnh1*Δ*, rrm3*Δ and $sen1^{cl}$ (a conditional lethal strain GAL::URL-HA-Sen1, shows lethality in glucose). (d) Meta-gene profile for RNA:DNA hybrid in G1 synchronized cells in wild type and *top2-1* cells. (e) Meta-gene profile for RNA:DNA hybrid comparison in wild type and *top2-1 top1*Δ double mutant. (f) Meta-gene profiles for RNA:DNA hybrid comparison in wild type, *hmo1*Δ, and *hmo1*Δ*top2-1* cells. (g) Gene density plot comparison of RNA:DNA hybrid comparison in wild type, *hmo1*Δ, and *hmo1*Δ*top2-1* cells. (h) Density plot showing the base coverage of RNA:DNA hybrids in genes with 'head-on' or 'co-directional' orientation with replication fork. Genes within 1 kb (upper panel n=347 genes) or 2 kb (lower panel n=539 genes) from the replication origins were considered. (i) Base coverage percentage of RNA:DNA hybrids accumulation at different intergenic space (<250 bp = 1729 gene pairs, 251-500bp = 2224 gene pairs and >500 bp = 2010 gene pairs) with respect to gene pairs grouped according to orientation.

Extended Data Figure 5. Ablation of negative supercoil at boundaries causes topological alterations and nucleosome repositioning
(a) Genome browser profile of chromosome-III from region 90 to 96 kb, showing comparative bTMP binding in wild type [Control], wild type [TopA], *top1*Δ *top2-1*[Control] and *top1*Δ *top2-1*[TopA] at 120 min at restrictive temperature. (b) Genome browser profile of chromosome V showing Hmo1 accumulation in wild type, *top2-1top1*Δ [Control], and *top2-1top1*Δ [TopA] after 120 min at restrictive temperature. (c) Absolute nucleosome count was derived from histone H3 ChIP-seq analysis in wild type [Control], wild type [TopA], *top2-1top1*Δ [Control] and *top2-1top1*Δ [TopA] at 120 min at restrictive temperature. (d) Nucleosome positions (150bp intervals) in wild type and *top2-1top1*Δ double mutants harboring either control or TopA-expressing plasmids were mapped with respect to TSS against Pol II gene percentage in y-axis.

Extended Data Figure 6. Top2 mediates chromatin-chromatin interactions
(a) Heat map showing PET clusters in all the chromosomes at 10 KB resolution. Inlet shows enlarged view of chromosome II. (b) Table showing the number of paired end reads, reads flanking with linker on both ends, unique mapped PETs and number of inter-ligation clusters obtained after Top2 ChIA-PET. (c) Density map showing PETs density at different interaction lengths. Inlet shows the enlarge view from 1000 to 10000 bp, which are considered for analysis. (d) Bar plot showing size wise distribution of chromatin loops. (e) Pie chart showing number of interactions matching with previous Top2-binding sites, number of interactions containing Top2 sites on either single or both side of the interaction, number of genes within and outside the Top2-mediated chromatin loops and number of loops with single gene or multiple genes inside. Other category includes genes partially overlapping with the loops and all other elements. (f) Genome browser view of ChIA-PET interactions along with Top2 and Hmo1 binding sites in chromosomes III and XVI

Extended Data Figure 7. A model for Hmo1, Top2 and Top1-mediated topological dynamics at Pol II genes.
(a) Schematic representation of cruciform DNA structures arising at negatively supercoiled DNA formed by two branches in a B-DNA duplex structure (red and black) and two branches in a non-B DNA duplex conformation (red or black). Cruciform DNA structures could form at gene boundaries and be stabilized by Hmo1. Gene looping is described in the blue area. (b) Schematic representation of gene loop structures in S-phase. Top2 associates with gene boundaries to harmonize topological transactions when transcribed genes are approached by incoming replication forks. We note that the topological dynamics described in the twin topological domain model[7] are not represented in our scheme. (c) In the absence of Hmo1, negative supercoil would lose the cruciform conformation and become a substrate for Top2. Top2 defects would delocalize Top1 at gene boundaries. The cruciform structures stabilized by Hmo1 would then become substrates for unscheduled Top1 activity that would convert them into aberrant intermediates such as ssDNA, nicks and knots. In *hmo1top2* double mutants Top1 is not recruited at the gene boundaries which remain in a negative supercoil context.

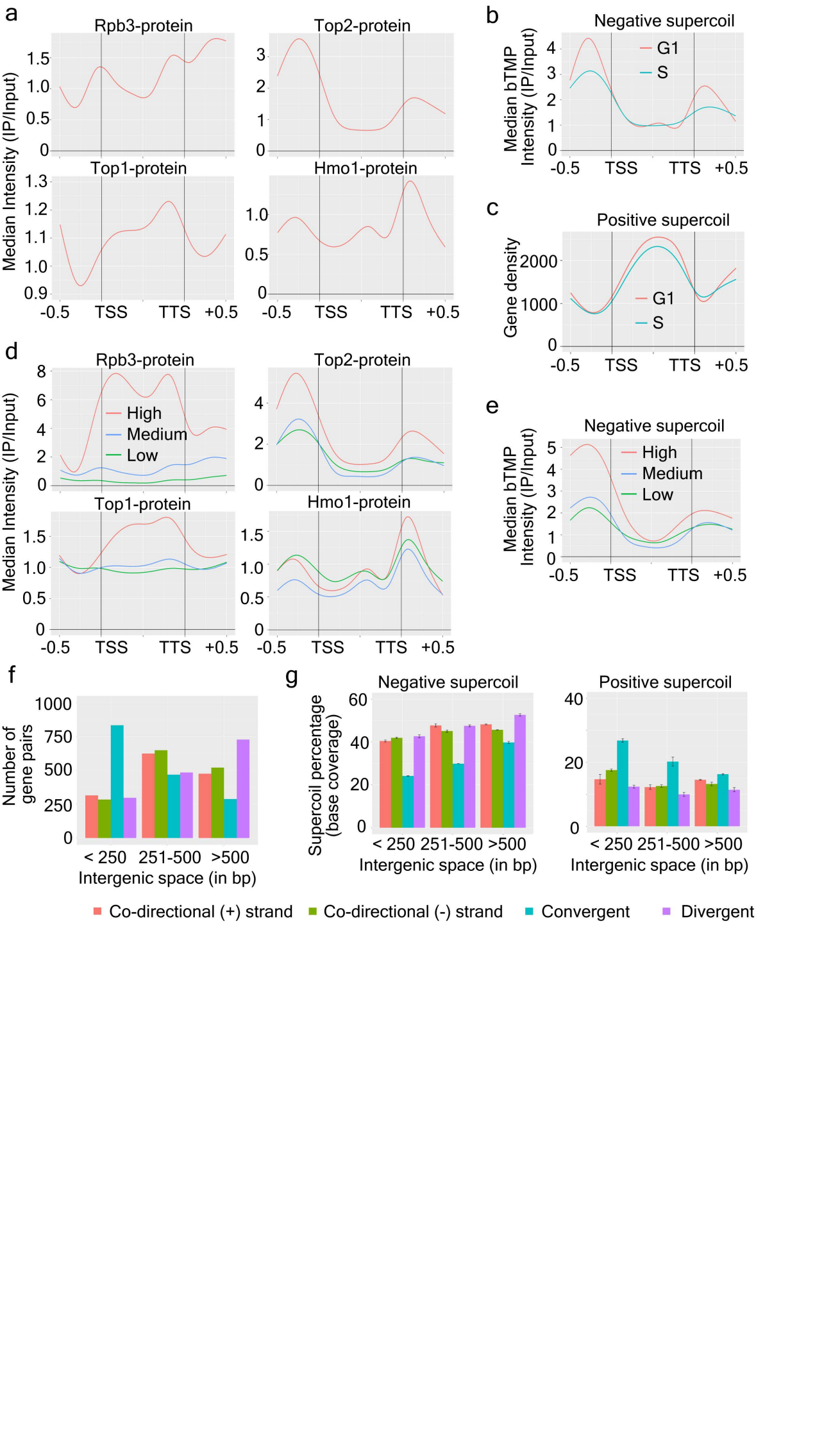
Extended Data Figure 8. bTMP titration
(a) By using biotin labelled oligo as a reference point, we titrated in vivo binding of bTMP in yeast at different concentration (from 0 to 800 ng). 4.16 pmol of biotin-oligos were diluted into 200 ul of TE. Respective microliter of oligo's were spotted on hybond membrane after equilibration with 1x maleic acid. Genomic DNA was isolated after UV (365 nM) crosslinking with respective amount of bTMP and ~300 ng (panel a) or 1500 ng (panel b) were spotted on the blot. The last sample was kept as negative control for UV crosslinking. In panel b, 5-fold more oligo and genomic DNA were spotted compared to panel a. Dot blot was developed with ExtrAvidin-Peroxidase antibody (Sigma-Aldrich cat no: E2886). (b) By measuring dot blot intensities, we estimated that 400 μg of bTMP was needed for $2x10^9$ yeast cells.

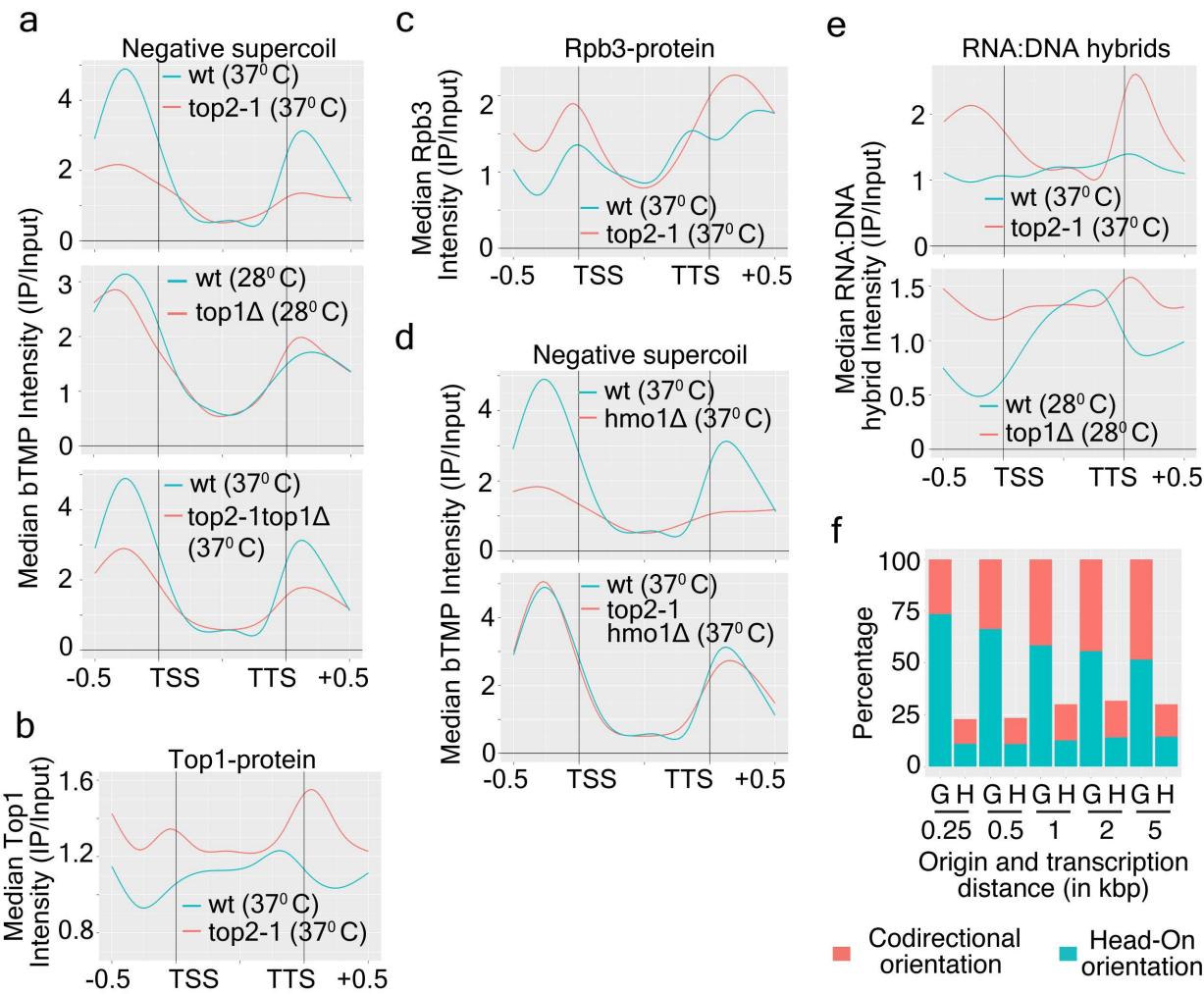
Extended Data Figure 9. bTMP binding normalization and dispersion profile for bTMP
(a) To avoid bias, a normalization was performed to filter potential sequence-specific psoralen DNA binding, as discussed in[23]. To normalize the data, we first purified and sonicated genomic DNA and then performed the bTMP-ChIP procedure on purified DNA fragments. The correction for microarray readings was done by subtracting bTMP binding *in vivo* with bTMP binding on purified/sonicated genomic DNA: (bTMP cells-IP/input) - (bTMP purified DNA-IP/input), which will give the normalized ratio of bTMP (bTMP-IP/input) binding. Meta ORF plot shows bTMP profiles with or without control DNA normalization of bTMP in wild type S-phase cells. (b) Meta-gene plot showing the normalized mean bTMP ratio and (c) median bTMP ratio in wild-type S-

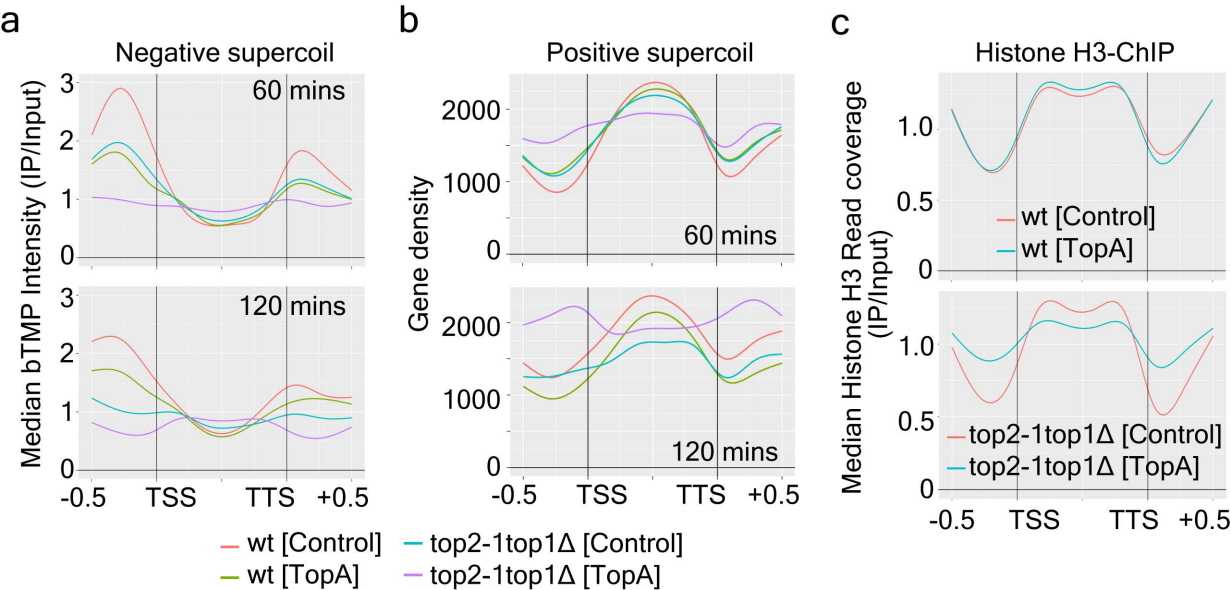
phase cells. For both the plots, bTMP binding ratios for all protein coding genes were plotted without smoothing, along with upper and lower confidence intervals (α = 0.05 or 95% limit;). Dotted line represents upper and lower confidence interval limit (α = 0.05 or 95% limit). The confidence interval does not deviate significantly from the mean and median values.

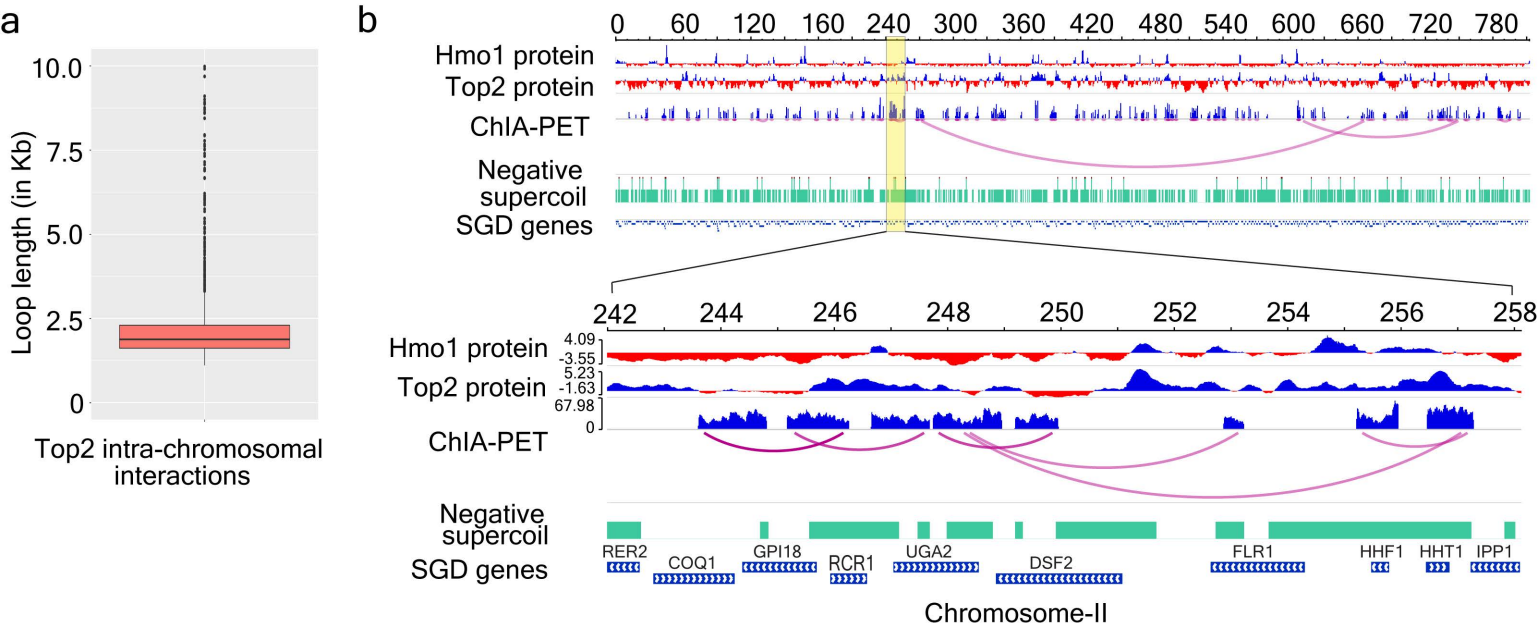
Extended Data Table 1. List of *S.cerevisiae* strains

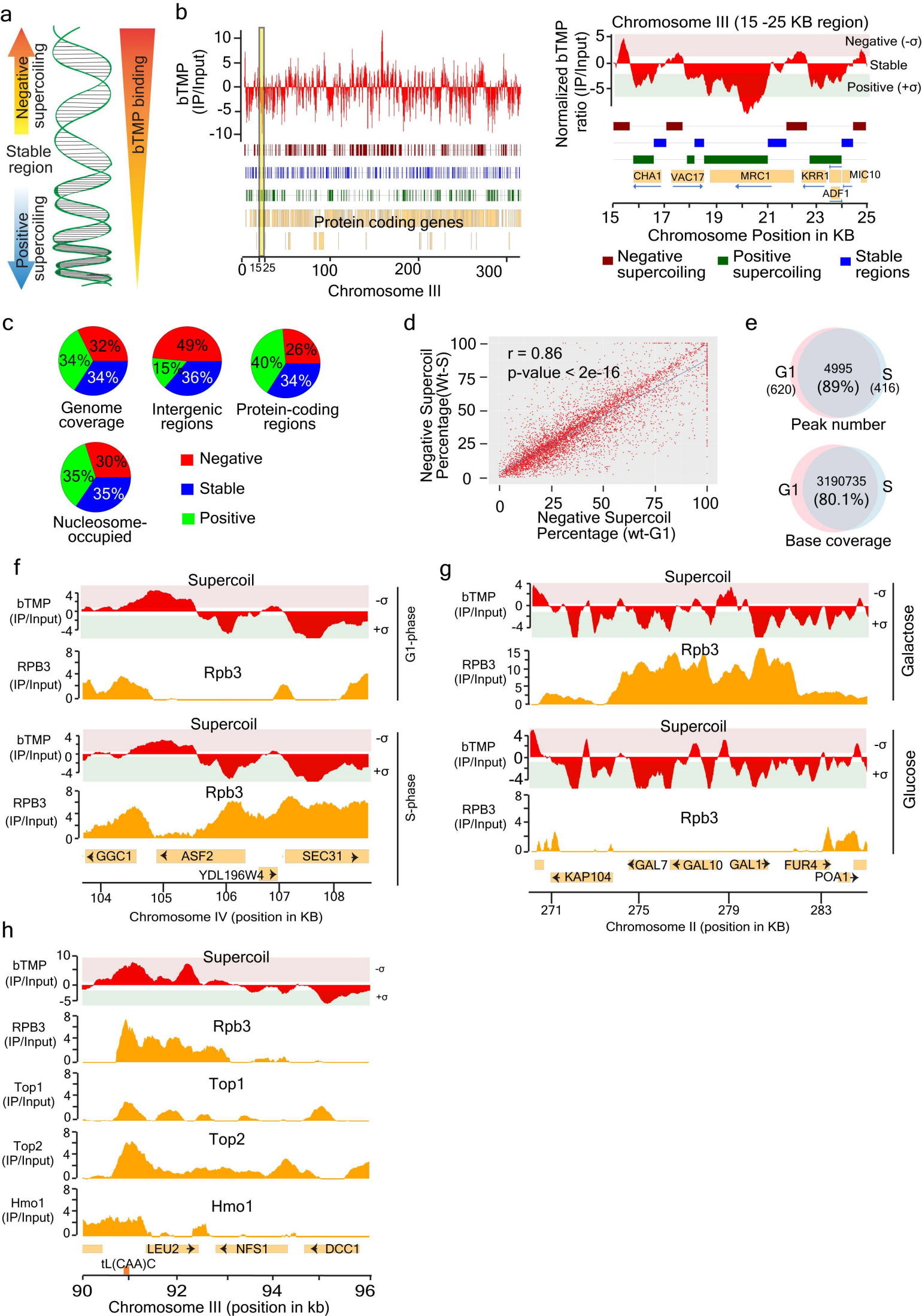
All *S.cerevisiae* strains used in this study along with relative genotypes are listed

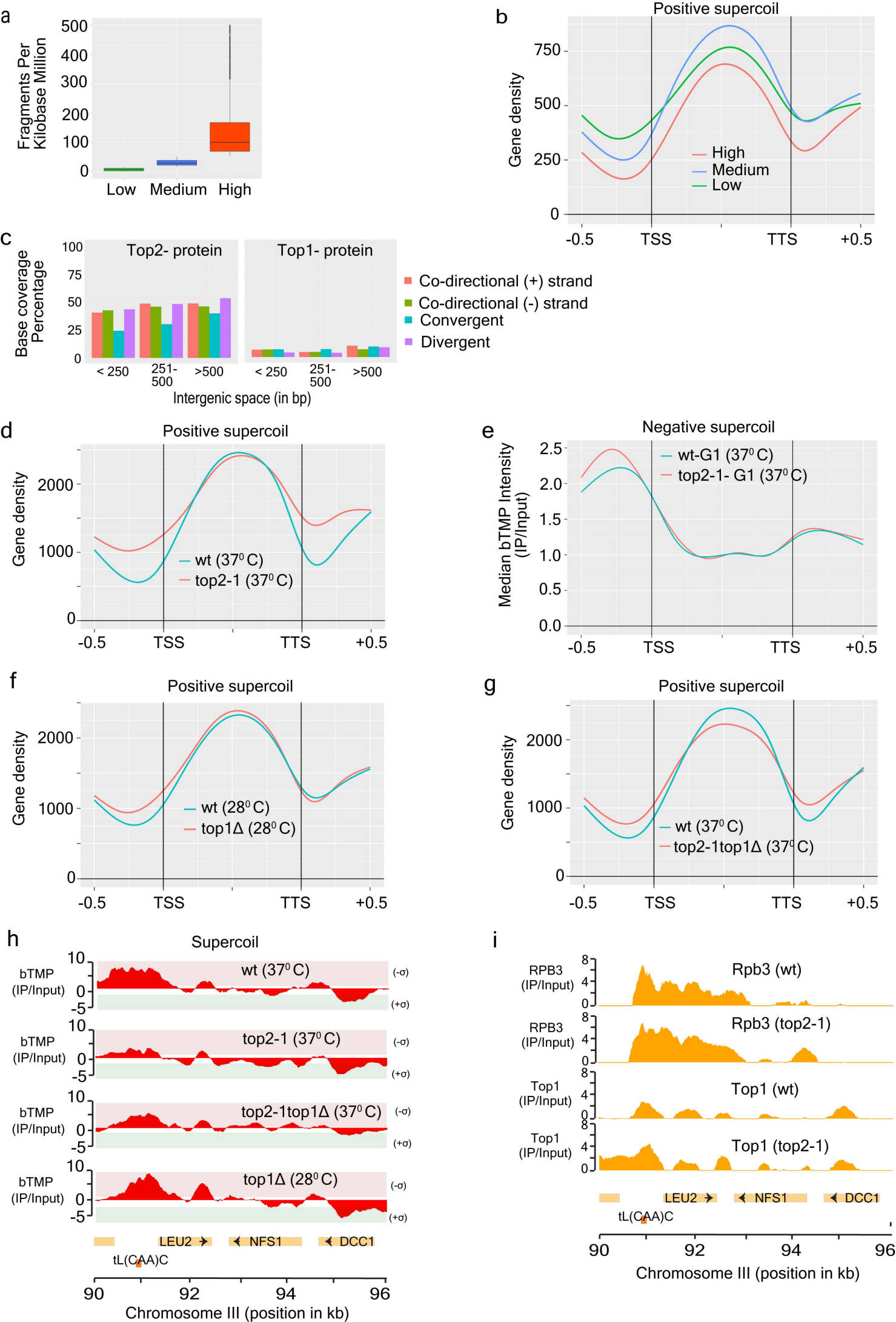

a
Rpb3-protein
Top2-protein
Top1-protein
Hmo1-protein
Median Intensity (IP/Input)
-0.5
TSS
TTS
+0.5
b
Negative supercoil
Median bTMP Intensity (IP/Input)
G1
S
c
Positive supercoil
Gene density
d
High
Medium
Low
e
f
Number of gene pairs
< 250
251-500
>500
Intergenic space (in bp)
g
Supercoil percentage (base coverage)
Co-directional (+) strand
Co-directional (-) strand
Convergent
Divergent

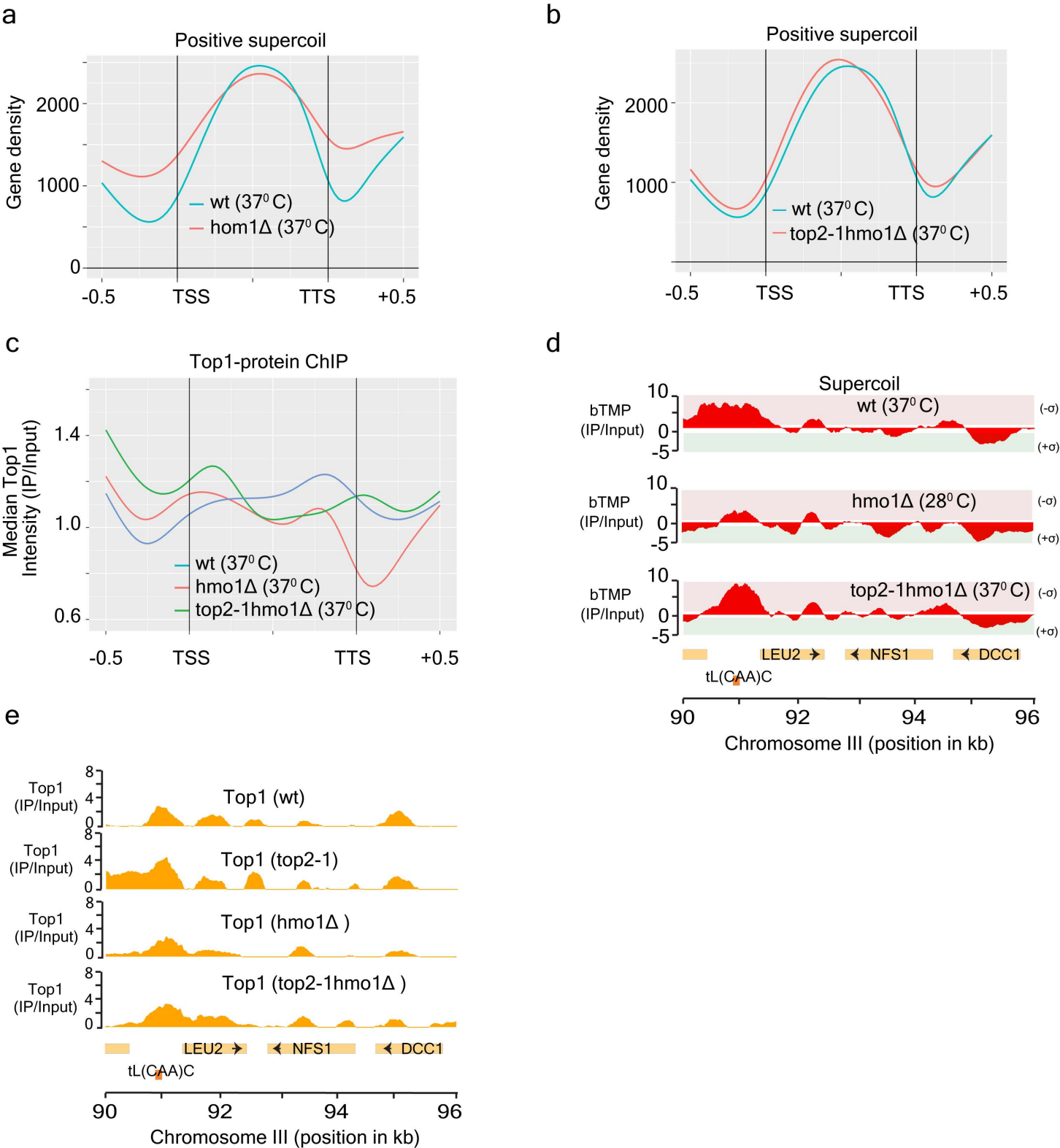

a
Negative supercoil
wt (37⁰ C)
top2-1 (37⁰ C)
wt (28⁰ C)
top1Δ (28⁰ C)
wt (37⁰ C)
top2-1top1Δ (37⁰ C)
Median bTMP Intensity (IP/Input)
-0.5
TSS
TTS
+0.5
b
Top1-protein
Median Top1 Intensity (IP/Input)
wt (37⁰ C)
top2-1 (37⁰ C)
c
Rpb3-protein
Median Rpb3 Intensity (IP/Input)
d
Negative supercoil
wt (37⁰ C)
hmo1Δ (37⁰ C)
top2-1 hmo1Δ (37⁰ C)
e
RNA:DNA hybrids
Median RNA:DNA hybrid Intensity (IP/Input)
f
Percentage
G H
0.25
0.5
1
2
5
Origin and transcription distance (in kbp)
Codirectional orientation
Head-On orientation

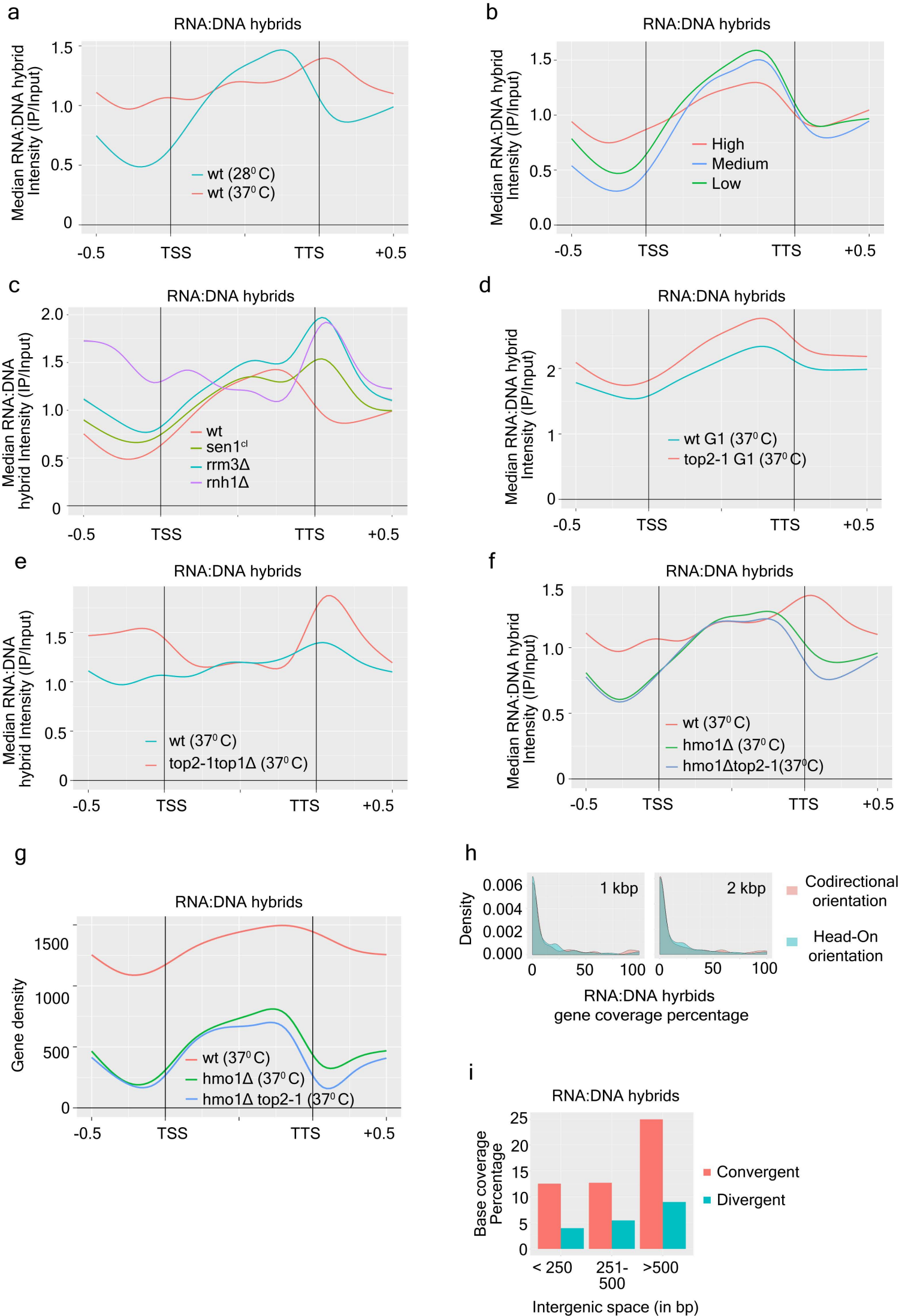

a
Negative supercoil
60 mins
120 mins
Median bTMP Intensity (IP/Input)
3
2
1
0
-0.5
TSS
TTS
+0.5
b
Positive supercoil
Gene density
2000
1000
0
60 mins
120 mins
c
Histone H3-ChIP
Median Histone H3 Read coverage (IP/Input)
1.0
0.5
wt [Control]
wt [TopA]
top2-1top1Δ [Control]
top2-1top1Δ [TopA]

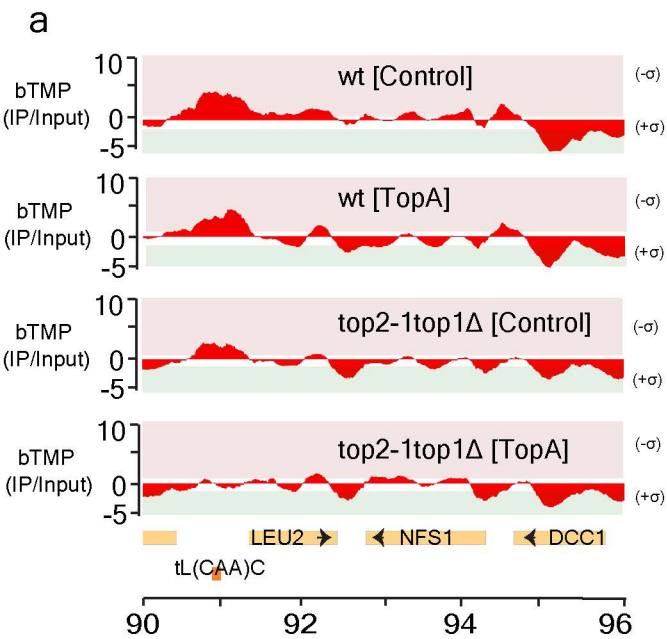

a
10.0
7.5
5.0
2.5
0
Loop length (in Kb)
Top2 intra-chromosomal interactions
b
0 60 120 160 240 300 360 420 480 540 600 660 720 780
Hmo1 protein
Top2 protein
ChIA-PET
Negative supercoil
SGD genes
242 244 246 248 250 252 254 256 258
Hmo1 protein
4.09
-3.55
Top2 protein
5.23
-1.63
ChIA-PET
67.98
0
Negative supercoil
SGD genes
RER2
COQ1
GPI18
RCR1
UGA2
DSF2
FLR1
HHF1
HHT1
IPP1
Chromosome-II

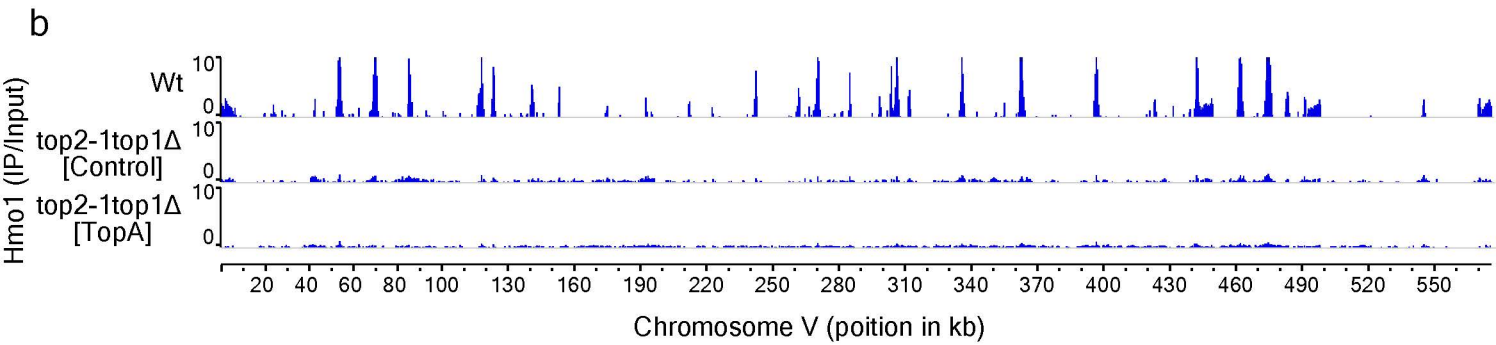

a
Negative supercoiling
Stable region
Positive supercoiling
bTMP binding
b
bTMP (IP/Input)
Protein coding genes
Chromosome III
Chromosome III (15 -25 KB region)
Normalized bTMP ratio (IP/Input)
Negative (-σ)
Stable
Positive (+σ)
CHA1
VAC17
MRC1
KRR1
MIC10
ADF1
Chromosome Position in KB
Negative supercoiling
Positive supercoiling
Stable regions
c
Genome coverage
Intergenic regions
Protein-coding regions
Nucleosome-occupied
Negative
Stable
Positive
d
Negative Supercoil Percentage(Wt-S)
r = 0.86
p-value < 2e-16
Negative Supercoil Percentage (wt-G1)
e
G1 (620)
4995 (89%)
S (416)
Peak number
G1
3190735 (80.1%)
S
Base coverage
f
Supercoil
bTMP (IP/Input)
RPB3 (IP/Input)
Rpb3
G1-phase
S-phase
GGC1
ASF2
SEC31
YDL196W4
Chromosome IV (position in KB)
g
Galactose
Glucose
KAP104
GAL7
GAL10
GAL1
FUR4
POA1
Chromosome II (position in KB)
h
Top1
Top2
Hmo1
LEU2
NFS1
DCC1
tL(CAA)C
Chromosome III (position in kb)

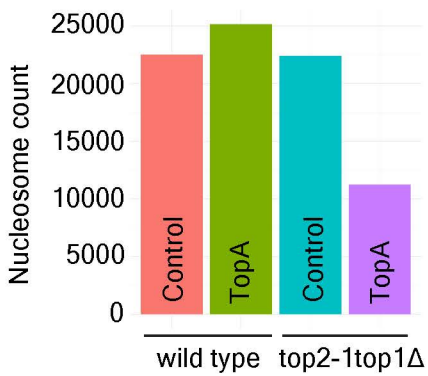

a
Fragments Per Kilobase Million
500
400
300
200
100
0
Low
Medium
High
b
Positive supercoil
Gene density
750
500
250
0
High
Medium
Low
-0.5
TSS
TTS
+0.5
c
Base coverage Percentage
100
75
50
25
0
Top2- protein
Top1- protein
< 250
251-500
>500
Intergenic space (in bp)
Co-directional (+) strand
Co-directional (-) strand
Convergent
Divergent
d
Positive supercoil
Gene density
2000
1000
0
wt (37⁰ C)
top2-1 (37⁰ C)
e
Negative supercoil
Median bTMP Intensity (IP/Input)
2.5
2.0
1.5
1.0
0.5
0.0
wt-G1 (37⁰ C)
top2-1- G1 (37⁰ C)
f
Positive supercoil
wt (28⁰ C)
top1Δ (28⁰ C)
g
Positive supercoil
wt (37⁰ C)
top2-1top1Δ (37⁰ C)
h
Supercoil
bTMP (IP/Input)
wt (37⁰ C)
top2-1 (37⁰ C)
top2-1top1Δ (37⁰ C)
top1Δ (28⁰ C)
(-σ)
(+σ)
LEU2
NFS1
DCC1
tL(CAA)C
90
92
94
96
Chromosome III (position in kb)
i
RPB3 (IP/Input)
Rpb3 (wt)
Rpb3 (top2-1)
Top1 (IP/Input)
Top1 (wt)
Top1 (top2-1)
LEU2
NFS1
DCC1
tL(CAA)C
90
92
94
96
Chromosome III (position in kb)

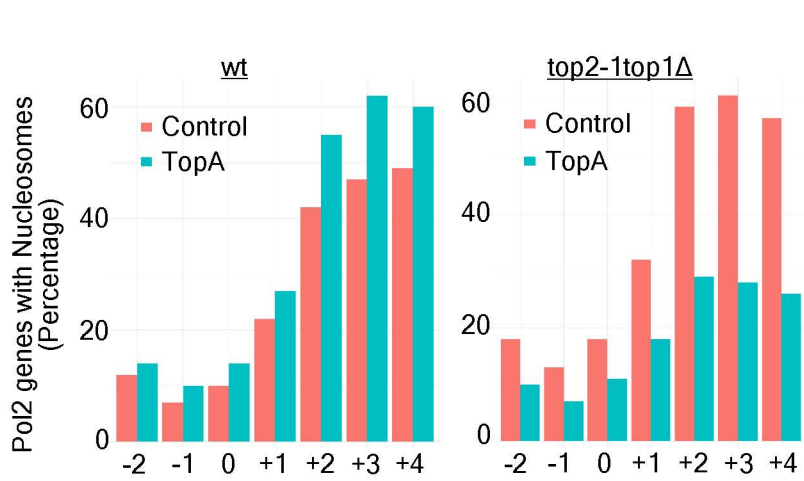

a
Positive supercoil
Gene density
2000
1000
0
wt (37⁰ C)
hom1Δ (37⁰ C)
-0.5
TSS
TTS
+0.5
b
Positive supercoil
Gene density
2000
1000
wt (37⁰ C)
top2-1hmo1Δ (37⁰ C)
-0.5
TSS
TTS
+0.5
c
Top1-protein ChIP
Median Top1 Intensity (IP/Input)
1.4
1.0
0.6
wt (37⁰ C)
hmo1Δ (37⁰ C)
top2-1hmo1Δ (37⁰ C)
-0.5
TSS
TTS
+0.5
d
Supercoil
bTMP (IP/Input)
10
0
-5
wt (37⁰ C)
hmo1Δ (28⁰ C)
top2-1hmo1Δ (37⁰ C)
(-σ)
(+σ)
LEU2
NFS1
DCC1
tL(CAA)C
90
92
94
96
Chromosome III (position in kb)
e
Top1 (IP/Input)
8
4
0
Top1 (wt)
Top1 (top2-1)
Top1 (hmo1Δ)
Top1 (top2-1hmo1Δ)
LEU2
NFS1
DCC1
tL(CAA)C
90
92
94
96
Chromosome III (position in kb)


a
RNA:DNA hybrids
Median RNA:DNA hybrid Intensity (IP/Input)
wt (28⁰ C)
wt (37⁰ C)
-0.5
TSS
TTS
+0.5
b
RNA:DNA hybrids
High
Medium
Low
c
RNA:DNA hybrids
wt
sen1cl
rrm3Δ
rnh1Δ
d
RNA:DNA hybrids
wt G1 (37⁰ C)
top2-1 G1 (37⁰ C)
e
RNA:DNA hybrids
wt (37⁰ C)
top2-1top1Δ (37⁰ C)
f
RNA:DNA hybrids
wt (37⁰ C)
hmo1Δ (37⁰ C)
hmo1Δtop2-1(37⁰C)
g
RNA:DNA hybrids
Gene density
wt (37⁰ C)
hmo1Δ (37⁰ C)
hmo1Δ top2-1 (37⁰ C)
h
Density
1 kbp
2 kbp
Codirectional orientation
Head-On orientation
RNA:DNA hyrbids
gene coverage percentage
i
RNA:DNA hybrids
Base coverage Percentage
Convergent
Divergent
< 250
251-500
>500
Intergenic space (in bp)

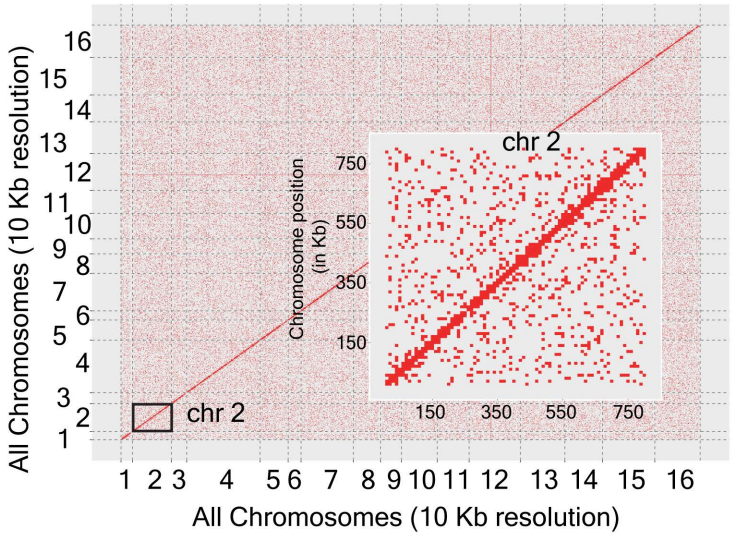

a
bTMP (IP/Input)
wt [Control]
wt [TopA]
top2-1top1Δ [Control]
top2-1top1Δ [TopA]
10
0
-5
(-σ)
(+σ)
LEU2
NFS1
DCC1
tL(CAA)C
90
92
94
96


b
Hmo1 (IP/Input)
Wt
top2-1top1Δ [Control]
top2-1top1Δ [TopA]
10
0
20 40 60 80 100 130 160 190 220 250 280 310 340 370 400 430 460 490 520 550
Chromosome V (poition in kb)


c
Nucleosome count
25000
20000
15000
10000
5000
0
Control
TopA
Control
TopA
wild type
top2-1top1Δ


d
wt
top2-1top1Δ
Control
TopA
Pol2 genes with Nucleosomes (Percentage)
60
40
20
0
-2 -1 0 +1 +2 +3 +4
Nucleosome position (150bp intervals) with respect to TSS

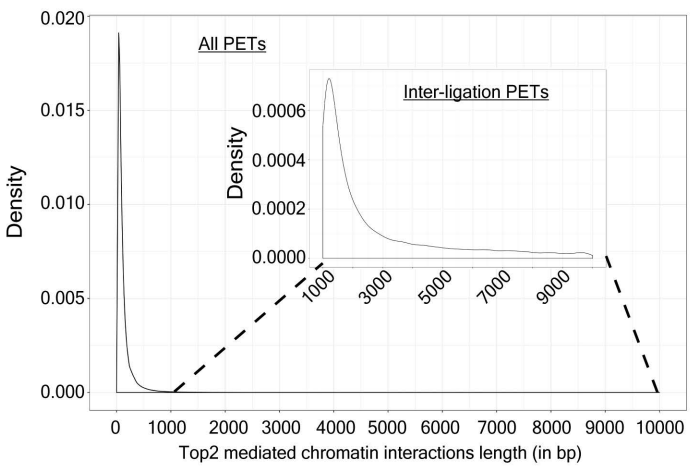
a



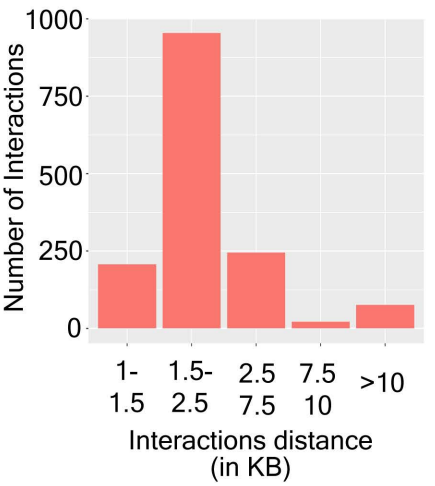
b

Total number of Paired End Reads	10586483
Reads flanking the linker on both ends	2868633
Unique mapped PETs	1238009
Inter-Ligation PET clusters (1kb Minimum Distance)	1887

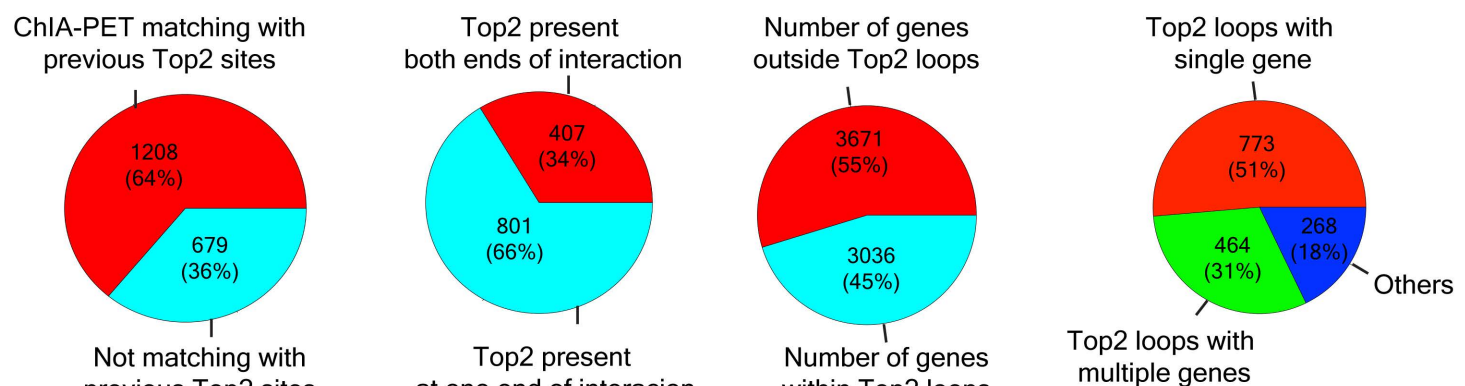
c



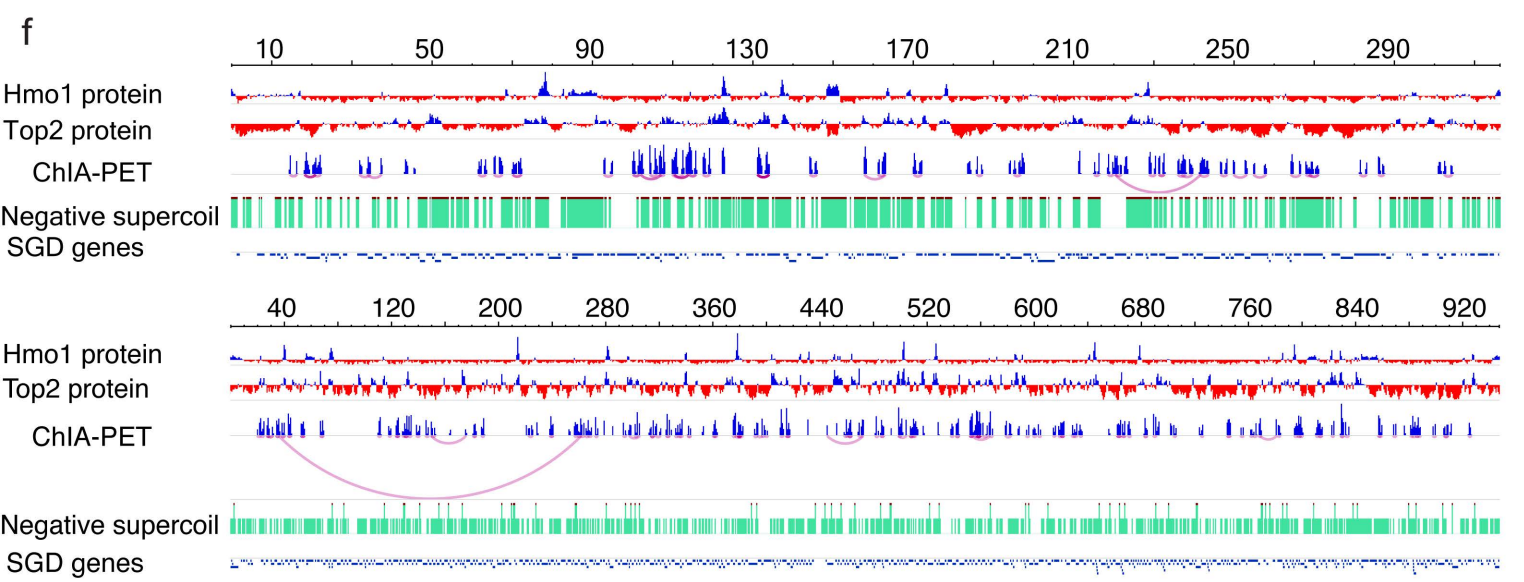
d

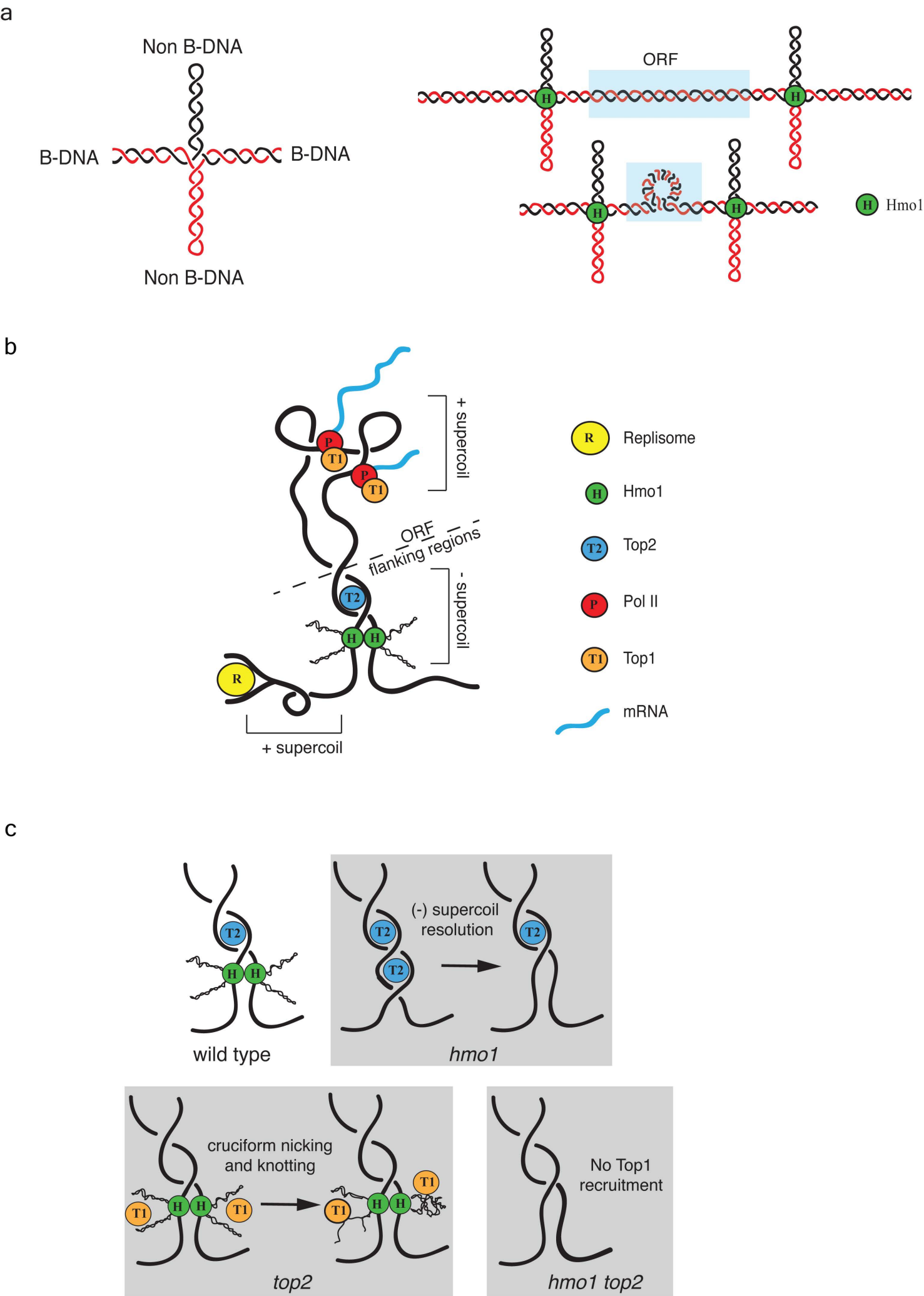


e

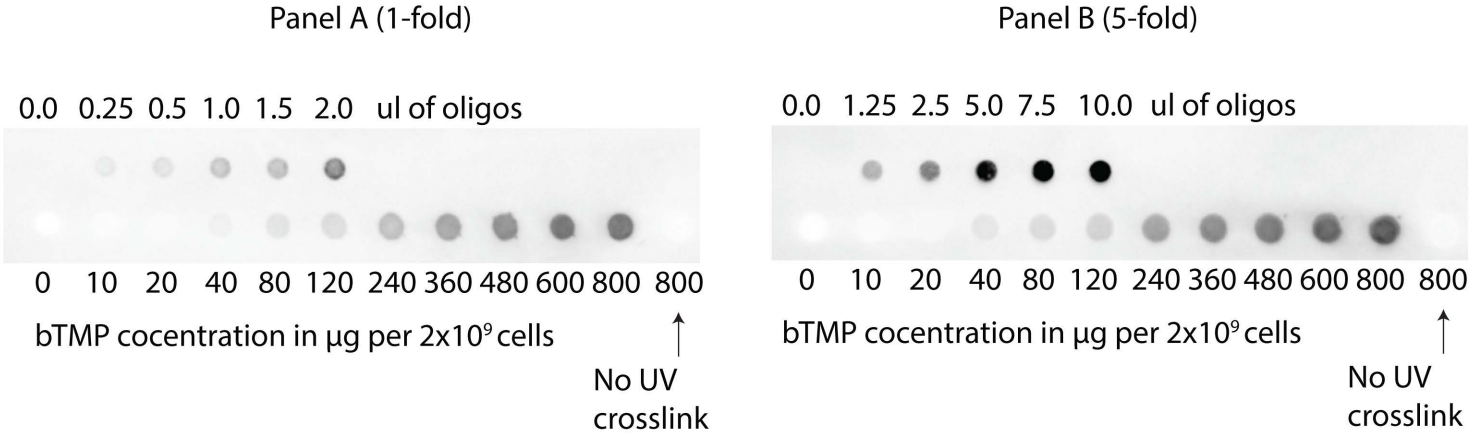


f

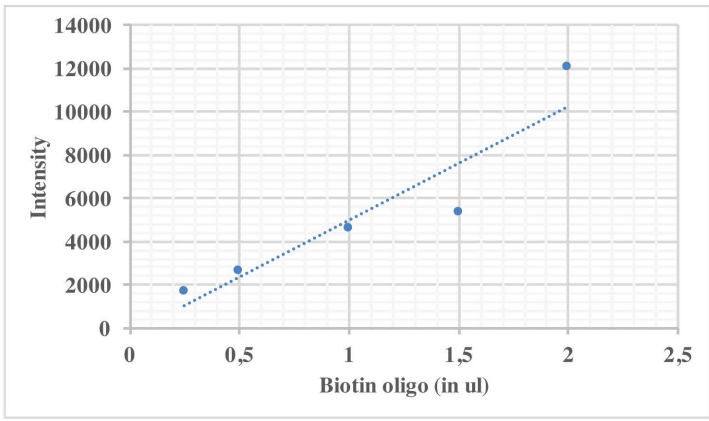


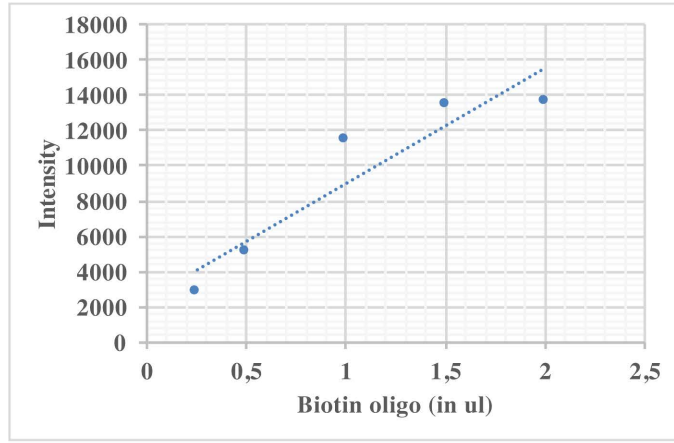

a
Non B-DNA
B-DNA
B-DNA
Non B-DNA
ORF
H
Hmo1
b
+ supercoil
P
T1
P
T1
ORF
flanking regions
- supercoil
T2
H
H
R
+ supercoil
R
Replisome
H
Hmo1
T2
Top2
P
Pol II
T1
Top1
mRNA
c
T2
H
H
wild type
(-) supercoil
resolution
T2
T2
T2
hmo1
cruciform nicking
and knotting
T1
H
H
T1
T1
T1
H
H
top2
No Top1
recruitment
hmo1 top2

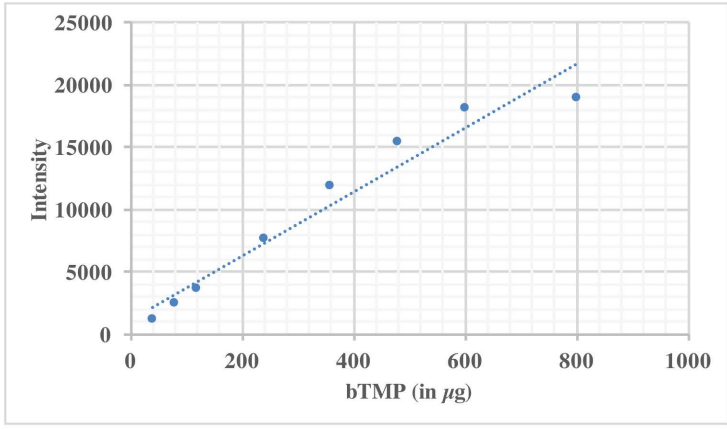
a

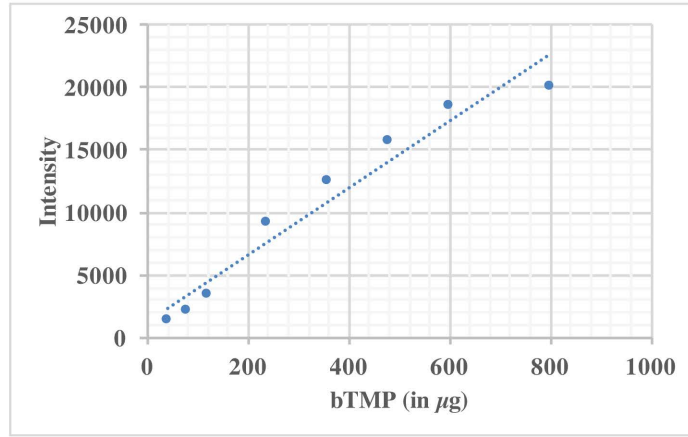

Panel A (1-fold)
0.0 0.25 0.5 1.0 1.5 2.0 ul of oligos
0 10 20 40 80 120 240 360 480 600 800 800
bTMP cocentration in µg per 2x10⁹ cells
No UV crosslink
Panel B (5-fold)
0.0 1.25 2.5 5.0 7.5 10.0 ul of oligos
0 10 20 40 80 120 240 360 480 600 800 800
bTMP cocentration in µg per 2x10⁹ cells
No UV crosslink

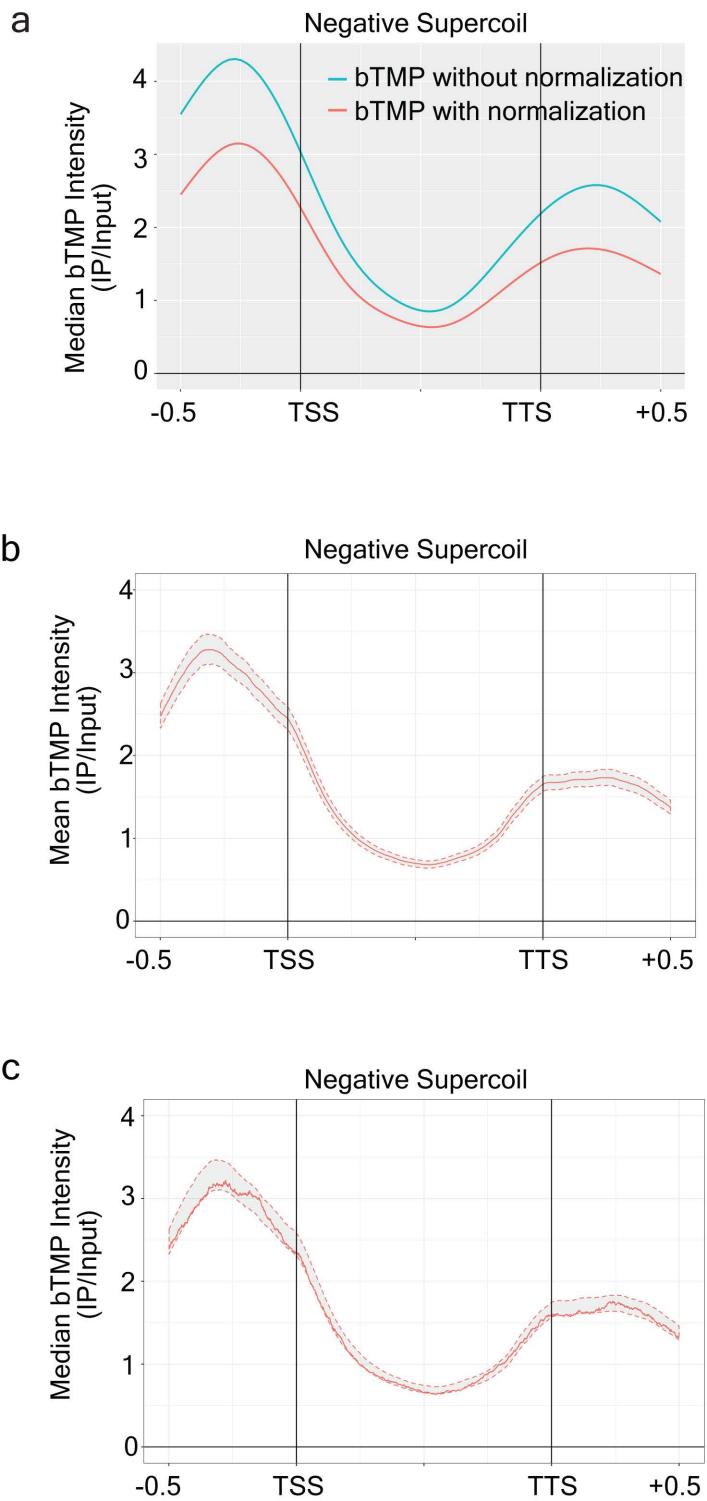
b


Intensity
Biotin oligo (in ul)


Intensity
Biotin oligo (in ul)


Intensity
bTMP (in µg)


Intensity
bTMP (in µg)


a
Negative Supercoil
bTMP without normalization
bTMP with normalization
Median bTMP Intensity
(IP/Input)
4
3
2
1
0
-0.5
TSS
TTS
+0.5
b
Negative Supercoil
Mean bTMP Intensity
(IP/Input)
4
3
2
1
0
-0.5
TSS
TTS
+0.5
c
Negative Supercoil
Median bTMP Intensity
(IP/Input)
4
3
2
1
0
-0.5
TSS
TTS
+0.5

	Strain	Stock Number	Genotype	Reference
1	Wt	SY2080	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+	Lab collection
2	*top2-1*	CY8423	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1	Lab collection
3	*hmo1*△	CY8476	MATa, ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+ ,hmo1::HIS	Lab collection
4	*top2-1hmo1*△	CY8475	MATa, ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 112 trp1-1, RAD5+, hmo1::HIS, top2-1	Lab collection
5	*top1*△	CY9950	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top1::HIS	Lab collection
6	Top1-6XHis-10xFlag	CY7178	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+, ura3::URA3/GPD-TK(7X), top1-6His10Flag (KANr)	Lab collection
7	*top2-1,*Top1-6XHis-10xFlag	CY7411	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+, ura3::URA3/GPD-TK(7X), top1-6His10Flag (KANr), top2-1	Lab collection
8	*top2-1top1*△	CY10344	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1, top1::HIS	Lab collection
9	*hmo1*△*,*Top1-6XHis-10xFlag	CY15215	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+, hmo1::Hygromycin, top1-6His10Flag (KANr)	This Study
10	*hmo1*△*top2-1,*Top1-6XHis-10x Flag	CY15216	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+, top1-6His10Flag (KANr), top2-1, hmo1::Hygromycin	This Study
11	Rpb3-10X-Flag	CY15214	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+, HIS3::BrdU-Inc, rpb3::RPB3-10X Flag-KanMX6	This Study
12	Wt [control]	CY15421	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+ [pYEp13-LEU empty]	This Study
13	Wt [TopA]	CY15422	Mata, ade2-1, ura3-1, trp1-1, leu2-3, leu2-112, his3-11, his3-15, can1-100, GAL, PSI+, RAD5+ [pJRW13-YEptopA-pGPD-LEU]	This Study
14	*top2-1top1*△ [control]	CY15423	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1, top1::HIS [pYEp13-LEU empty]	This Study
15	*top2-1top1*△ [TopA]	CY15424	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1, top1::HIS [pJRW13-YEptopA-pGPD-LEU]	This Study
16	*top2-1top1*△*,*Hmo1-10X Flag, [Control]	CY15427	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1, top1::HIS, HMO1::HMO1-10X Flag [pYEp13-LEU empty]	This Study
17	*top2-1top1*△*,*Hmo1-10X Flag, [TopA]	CY15428	MATa ADE2+ CAN1+, ura3-1, his3-11,15 leu2-3, 12 trp1-1, RAD5+, top2-1, top1::HIS, HMO1::HMO1-10X Flag [pJRW13-YEptopA-pGPD-LEU]	This Study